Journal of the
OPTICAL SOCIETY
of AMERICA

VOLUME 39, NUMBER 3 MARCH, 1949

Low Temperature Radiation Pyrometry in Industry

J. C. MOUZON AND C. A. DYER
The Brown Instrument Company, Division of Minneapolis-Honeywell Regulator Company, Philadelphia, Pennsylvania

(Received October 11, 1948)

Problems encountered in the application of radiation pyrometry in industry are discussed with particular emphasis being placed on the measurement of low surface temperatures (100–400°F). A low range radiation pyrometer for use on such applications is described.

INTRODUCTION

FOR a long time there has been a demand for improved means for measuring the temperature of moving objects in the range 100°F to 1000°F. This demand exists in the fields of rubber, plastics, and paper processing; panel heating, induction and dielectric heating, and infra-red drying. The purpose of this paper is to discuss some of the problems involved in low temperature radiation pyrometry and to describe an industrial radiation pyrometer which fills a need of long standing.

GENERAL CONSIDERATIONS

The curves of Fig. 1 show the spectral energy distribution from a blackbody over the range of temperature 32°F to 441°F. An examination of these curves will familiarize one with the wave-lengths associated with radiation measurements in this range of temperatures. It is seen that the peak of the emission curve corresponding to a temperature of 441°F lies at about 6 microns. (It is well to keep in mind that only 25 percent of the energy under a blackbody curve lies to the left of the peak of the curve.) For convenience, the visible light region is shown, and the approximate cut-off points of Pyrex, fused silica, and fluorite lenses are indicated. The superior transmission of fluorite for radiant energy of these wave-lengths is obvious.

RADIATION DETECTORS

Suitable detectors of energy of the wave-lengths in question are rather limited. In general, one may list: thermocouples, bolometers, photo-cells, pneumatic heat detectors, infra-red phosphors, and vane type radiometers. Neither the infra-red phosphors nor the lead sulfide photo-cells are sensitive far enough into the infra-red for our purposes; the phosphors are good only in the very near infra-red, and the lead sulfide cell cuts off at approximately 3 microns unless cooled in liquid hydrogen.

Since an industrial instrument is a device which is expected to give trouble-free service under most adverse conditions, the field of suitable radiation detectors is narrowed still further. For the instrument herein described, a 10-junction thermopile was selected as the detecting element.

SURFACE EMISSIVITY

A rather serious problem in all radiation measurements is that associated with the emissivity of the surfaces involved. Here the term emissivity* specifies the ratio of the rate of energy emitted from a given body to the rate at which energy would be emitted from a blackbody at the same temperature.

Table I[1] lists the emissivities of various metal surfaces, both oxidized and unoxidized. The large variation from the unoxidized to the oxidized con-

* Note that no distinction is made here between the meanings of the words emittance and emissivity.

[1] Taken from the appendix of *Temperature, Its Measurement and Control in Science and Industry* (Reinhold Publishing Corporation, New York, 1941).

TABLE I. Total emissivity of metals.

Material	Surface unoxidized 25°C	100°C	500°C	Surface oxidized 200°C
Aluminum	0.022	0.028	0.060	0.11
Chromium		0.08		
Copper		0.02		0.18
Iron		0.05		0.74—100°C
Lead		0.05		0.63
Nickel	0.045	0.06	0.12	0.37
Steel		0.08		0.79
Monel				0.43

dition presents a very serious problem in radiation measurements of the temperature of metals.

Figure 2[2] presents some rather interesting data on copper. Note the change in emissivity as one adds lacquer to a bright copper surface. For those interested in low temperature measurements, the data in Fig. 3[3] are very comforting. Here it is shown that the emissivities of various paints and refractories approach a value of 90 percent or more at 200°F. Those accustomed to working in the infra-red recognize that the visible color of a surface is not related to its emissivity in the infra-red.

It is of prime importance for radiation measurements that the emissivity of the body in question stay constant if reproducible results are to be obtained.** It is further very desirable, if not essential, that the emissivity of the object whose temperature is to be measured be large. Existing data indicate that the emissivity of such things as plastics and paints is high and stays constant over long periods of time. A polished metal roll, however, may sometimes have a fairly high emissivity because of a very thin film covering it. Usually, such rolls have very low emissivities, as is suggested by the data in Table I. Furthermore, the emissivity of such rolls may change with time by a very large percent merely by the formation of a thin film of oxide or other material. It is, therefore, inadvisable as a general rule to attempt to measure the temperature of polished rolls by radiation means. A further reason for advising against the measurement of surface temperatures of low emissivity surfaces is that the total radiation falling upon the detector may to a large extent be made up of radiation reflected from the surface whose temperature is being measured. Thus there is danger of measuring primarily the temperature of the surroundings instead of the temperature of the desired surface. This difficulty can, however, be avoided in practice by proper shielding.

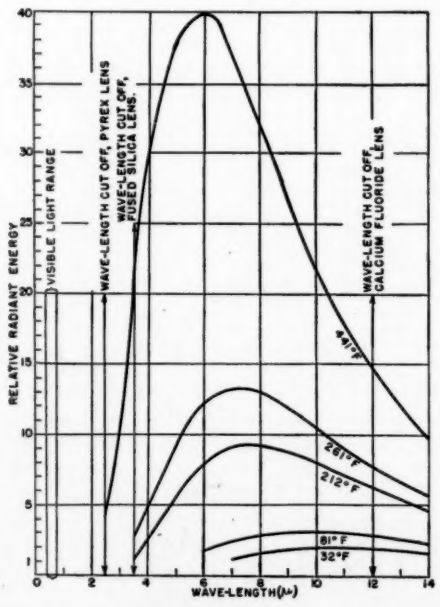

FIG. 1. Spectral energy distribution from a blackbody at various temperatures.

Such problems exist to some extent at high temperatures as well as at low temperatures. Attention has been given them in the measurement of high surface temperatures by radiation pyrometry. There are cases in industry where the emissivity correction is from 200°F to 600°F at surface temperatures of 1200°F. The problem of emissivities is not then unique with low temperature measurements.

The calibration of a low temperature radiation pyrometer is something which needs careful interpretation. It is perfectly true that users of such equipment would be very pleased to know precisely the temperature of the surface in question, but for industrial measurement and control, it is probably more satisfactory to make an empirical calibration of the equipment on the job than it is to attempt to use a blackbody calibration. Experience and the job at hand will dictate how important a blackbody calibration may be.

REFERENCE TEMPERATURES

The actual temperature of any detector of radiation is a function not only of the temperature of the object from which it receives radiation, but also the temperature of its own immediate surroundings. It receives radiation from its own support or housing as well as from the outside, and likewise loses energy by radiation and conduction.

[2] See reference 1, p. 1182.
[3] See reference 1, p. 1186.
** It is recognized that if the temperature of the detector is controlled at the measured temperature, under special conditions, the variation in emissivity of a body will not produce an error in the measurement, but will only change the sensitivity of the equipment. This scheme would be difficult to apply over a wide range of temperatures, but has considerable merit in principle.

Regardless of the nature of the detector used for radiation measurements, its output is very much dependent on the temperature of its immediate surroundings. The temperature of the surroundings of the detector is actually the reference temperature.

The sensitivity of the detector to changes in ambient temperature is in part dependent on the temperature of the source on which it is sighted. If the temperature of the source is high, the output of the detector is roughly (not exactly) proportional to the fourth power of the absolute temperature of the source. On the other hand, if the temperature of the source is near the temperature of the detector, the output of the detector is, to a crude approximation, proportional to the first power of the difference in the absolute temperature of the source and the housing in which the detector is enclosed.

During the war much work was done on problems of infra-red detection. However, most of the thinking was confined to the problem of detecting an object against a background having a different temperature. A change in the temperature of the housing of such a detector unit produced, at worst, only a gradual drift in the output of the detector. If the detector happened to be a bolometer with chopped radiation falling on it, and if the bolometer were capacitively coupled to an amplifier, normal changes in ambient temperature would produce no significant drift. Such detectors are unsuited for low temperature measurements unless the housing temperature is controlled, or unless some means is provided for compensating for ambient temperature changes.

WINDOWS AND LENSES

Industrial radiation pyrometers require some form of protection for the detector element. It, therefore, must be enclosed in a housing, and either a window or a lens must be provided which can be cleaned periodically of dust and dirt. The materials available for windows and lenses are rather limited. During the war, thin films were used on some of the infra-red detector units for windows. There are two objections to the use of thin films in the present application. First, the transmission of most thin films is very poor in the region of one to twelve microns, and, second, a thin film is entirely unsuited as a window material for an industrial instrument because of its fragility.

In Table II a list of the transmissions of various optical materials is shown. A rough estimate has been made of the useful range of each material, and an approximate value has been given for the transmission over the arbitrarily chosen useful range. The cut-off wave-length for each material

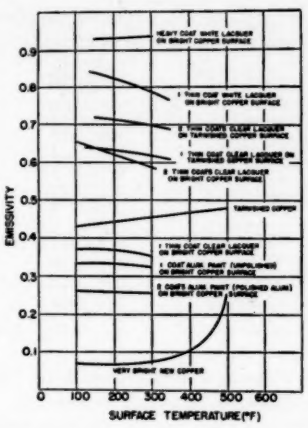

FIG. 2. Emissivity of copper under various surface conditions.

is also shown, as well as any prominent absorption bands that may exist. Reference to this table, as well as to Fig. 1, leads one almost immediately to the choice of fluorite as a suitable window or lens material for an industrial instrument. Rocksalt is considered out of the question. Silver chloride is also considered unsuited for this type of application in spite of the fact that reasonably good protective coatings have been developed. Thallium bromide-iodide is still unavailable. Fluorite has a melting point of 1330°C, and is so slightly soluble in water that the deterioration from water vapor is negligible. It has a hardness of 4 on Moh's scale, which means that it is hard enough to be cleaned without scratching. The fact that it has a very large coefficient of linear expansion is not particularly serious where moderate size pieces are used.

In Table III are shown data on the index of refraction of fluorite at various wave-lengths. In Fig. 4 is shown the transmission curve for the actual lens used in the instrument to be described. The

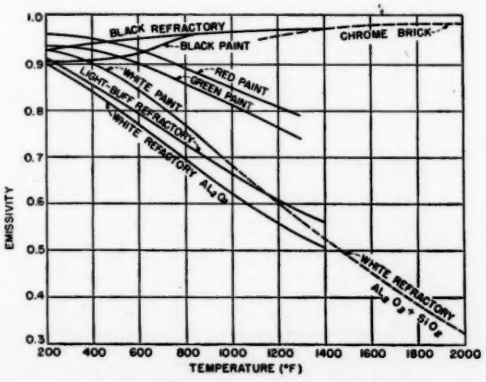

FIG. 3. Emissivities of various materials as a function of temperature.

TABLE II. Transmission ranges of various optical materials.

Optical material	Useful wave-length range, microns	Approximate transmission over useful range; %	Approximate cut-off wave-length, microns	Absorption bands, microns	Authority
Fluorite	0.2 −10	90	12		W. W. Coblentz
Glass (in general)	0.3 − 1	90	1		Corning Glass
Lithium fluoride	0.1 − 5	80	8		Harshaw Laboratories
Potassium bromide	0.2 −22	90	30		Harshaw Laboratories
Pyrex	0.35− 2.5	90	2.5		Corning Glass
Quartz	0.2 − 3	90	7	2.95	W. W. Coblentz
Rocksalt	0.2 −12	95	21		W. W. Coblentz
Silver chloride	1 −20	60–80	25		H. C. Kremers
Sylvite	0.3 −20	80	24	3.18, 7.08	W. W. Coblentz
Thallium bromide-iodide	2 −30	70	70		O. F. Tuttle and P. H. Egli

curve was obtained from standard transmission data on a sheet of a given thickness, corrections having been made for the variation in the thickness of the actual lens. In Fig. 5 is a curve showing the variation of the focal length of the lens with wave-length. There is considerable chromatic aberration with the lens used, as can be seen from the curve. This, however, is not too serious in the present instrument. In Fig. 6 the energy received by the thermopile from a blackbody at 212°F through the lens is compared to that received from the same body without the lens.

HUMIDITY EFFECTS

Water vapor absorption bands are plentiful in the region of 1 to 12 microns. Hence, radiation measurements in the infra-red are appreciably affected by humidity. The data in the literature on the absorption of water vapor as a function of path length are not too consistent, nor are they in such form as to make the evaluation of the absorption of total radiation passing through an absorbing lens a simple matter. Figure 7 shows some data on the variation of the output of the radiation pyrometer as a function of humidity and distance (25 g/m³ corresponds to a relative humidity of 99 percent at 80°F). Although the data are not as complete as might be desired, they are sufficient to indicate the magnitude of variations that can be expected when using a total radiation receiver.

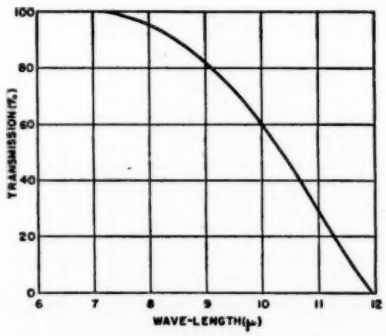

FIG. 4. Transmission of low temperature radiation pyrometer lens.

STABILITY OF THE RADIATION RECEIVER

It is interesting to investigate the effect of variations in the radiation receiver on the calibration of a radiation pyrometer. If the emissivity of a thermocouple receiver, for example, changes with time, or if the thermocouple changes its e.m.f. temperature characteristics with time, a corresponding error in calibration naturally results. A change in the emissivity of the receiver is a real problem and is more likely to occur than a change in thermocouple characteristics. It is shown below that such problems are more serious with low temperature radiation pyrometers than with those used at high temperatures.

Consider a radiation pyrometer consisting of a housing at temperature T_3, inside of which is a thermopile receiver at temperature T_2 which looks out at a blackbody at temperature T_1 through a perfectly transparent lens. Figure 8 shows a *small* source of area A_1 which sends radiation out through the cone whose half-angle is θ_1 into the lens and thence onto the *small* thermopile whose area is A_2. Let J be the power radiated from the blackbody per unit solid angle normal to the surface. The power radiated through the lens from the source at temperature T_1 is $A_1\pi J \sin^2\theta_1$.[4] If θ_1 were 90°, the power radiated would be $A_1\pi J$. This, however, is also equal to $\sigma A_1 T_1^4$ according to the Stefan-Boltzman Law. We may, therefore, write that the radiation received by the thermopile from the source is $e\sigma A_1 T_1^4 \sin^2\theta_1$, where e is the emissivity of the front side of the thermopile receiver. If the back side of the thermopile receiver is a perfect reflector, the heat received by the thermopile from the housing will be $e\sigma T_3^4 A_2(1-\sin^2\theta_2)$. The heat lost by the thermopile by radiation is $e\sigma A_2 T_2^4$, and the heat lost by conduction is $C(T_2-T_3)$. Thus (if

[4] J. M. Cork, *Heat* (John Wiley and Sons, Inc., New York, 1942), second edition, p. 148.

$A_1 \sin\theta_1 = A_2 \sin\theta_2$ for simplicity),

$$e\sigma T_1^4 A_2 \sin^2\theta_2 + e\sigma T_3^4 A_2(1-\sin^2\theta_2) = e\sigma T_2^4 A_2 + C(T_2 - T_3).$$

Rearranging and factoring,

$$T_1^4 - T_3^4 = (1/\sin^2\theta_2)[(T_2+T_3)(T_2^2+T_3^2) + (C/e\sigma A_2)](T_2 - T_3). \quad (1)$$

In many applications T_2 and T_3 are close enough together so that the variation in the right-hand side of Eq. (1) is largely determined by $T_2 - T_3$, while the term $(T_2+T_3)(T_2^2+T_3^2)$ remains relatively constant. With this approximation we can write:

$$T_1^4 - T_3^4 = (K_1 + (K_2/e))(T_2 - T_3). \quad (2)$$

The e.m.f. output of the thermopile is

$$E = K_3(T_2 - T_3) = (K_3/K_1 + (K_2/e))(T_1^4 - T_3^4).$$

For high temperatures $T_1^4 \gg T_3^4$, so

$$T_1^4 = (E/K_3)(K_1 + (K_2/e)),$$

$$T_1 = ((E/K_3)(K_1 + (K_2/e)))^{\frac{1}{4}}. \quad (3)$$

Equation (3) is the equation of an approximate calibration curve of temperature *vs.* thermopile output at high temperatures. Variations in K_3 or e are reflected as smaller variations in T_1 because of the 4th root.

For the measurement of low temperatures Eq. (2) can be written approximately as

$$T_1 - T_3 = (K_1' + (K_2'/e))(T_2 - T_3).$$

If the thermopile output is proportional to $T_2 - T_3$,

$$E = K_3'(T_2 - T_3) \text{ and}$$

$$T_1 - T_3 = (E/K_3')(K_1' + (K_2'/e))(T_2 - T_3).$$

$$T_1 = (E/K_3')(K_1' + (K_2'/e))(T_2 - T_3) + T_3. \quad (4)$$

TABLE III. Indices of refraction of fluorite in air at 20°C.

Wave-length microns	Index of refraction	Wave-length microns	Index of refraction
0.48615	1.43704	2.5537	1.42080
0.58758	1.43388	2.750	1.41956
0.65630	1.43249	3.0500	1.41750
0.76653	1.43093	3.5359	1.41376
1.0140	1.42884	4.000	1.40963
1.250	1.42752	4.6000	1.40357
1.5715	1.42596	5.000	1.39908
1.650	1.42558	5.8932	1.38712
1.8400	1.42468	6.4825	1.37824
2.0582	1.42360	7.6612	1.35675
2.250	1.42258	8.8398	1.33075

Equation (4) is the equation of an approximate calibration curve of temperature *vs.* thermopile output at low temperatures. Variations in K_3' or e are seen to be much more serious in this case than in the case of high temperature measurements. A much more stable detector is therefore required for low temperatures than for high temperatures. This requirement has been demonstrated in practice.

It might be noted from Eq. (4) that T_3 must be known accurately if T_1 is to be determined with precision. This is added proof that the housing temperature must be accurately controlled, or some means for compensation for housing temperature variation must be employed.

THE LOW RANGE RADIATION PYROMETER

Figure 9 is a sectional view of a radiation pyrometer developed for use in making low temperature measurements. It consists of two principal assemblies. One of these is the outer housing, upon which are located the lens and a heater winding. The other smaller assembly fits inside the first and

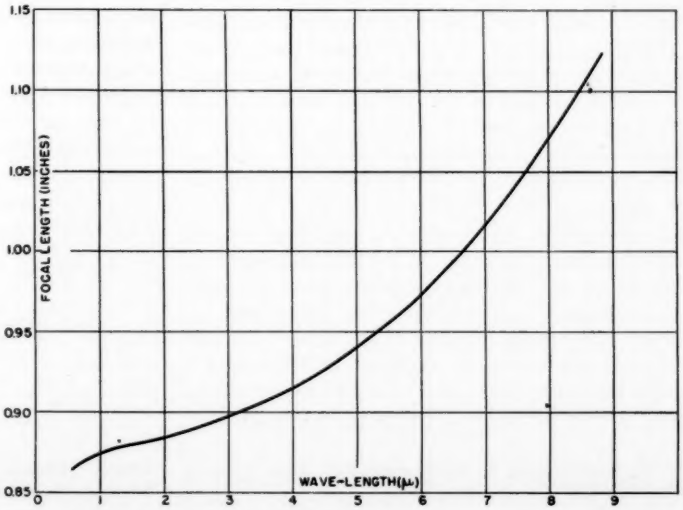

FIG. 5. Variation of focal length of a fluorite lens with wave-length. (Radii of curvature $R_1 = R_2 = 0.75$ inch.)

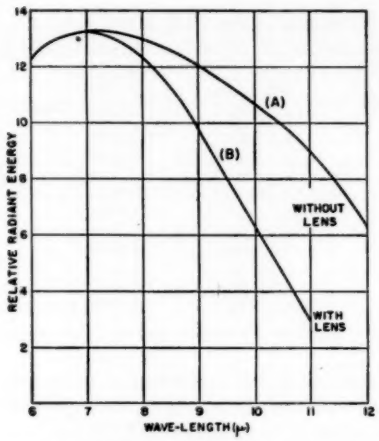

FIG. 6. Relative energy received by thermopile from a blackbody at 212°F.

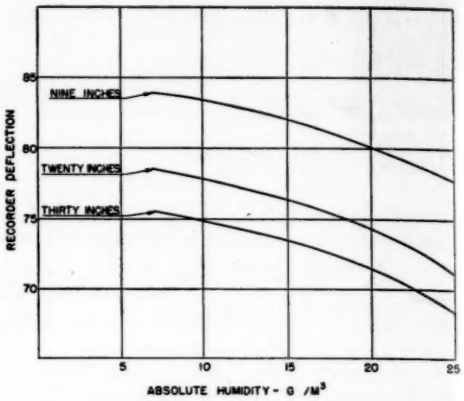

FIG. 7. Radiation pyrometer output as a function of humidity.

consists of a small housing upon which the thermopile, a small concave mirror, and a resistance thermometer winding are mounted. Electrical connections to the thermopile, the resistance thermometer, and heater windings are made to solder lugs mounted on a terminal disk at the rear of the assembly, but not shown in Fig. 9. The inner or thermopile housing is attached to this terminal disk.

As mentioned previously, the radiation detector is a ten-junction thermopile[5] assembled in such a manner that the hot junctions lie in a single plane and together form a receiving surface about $\frac{1}{8}''$ in diameter. The thermopile cold junctions are arranged so that they are held in intimate thermal contact with the pyrometer housing. A small concave mirror located directly behind the thermopile hot junctions serves two purposes. One of these is to provide some extra output from the thermopile by reflecting radiation onto the back of the thermopile that normally passes through its openings. The second purpose of the mirror is to provide an adjustment in thermopile output so that all pyrometers may be made interchangeable.

The lens chosen for the pyrometer is ground from artificially grown calcium fluoride crystal. It has been recessed well back into the main housing under the heater winding in an effort to keep its temperature as close to the thermopile cold junction temperature as possible. The thermopile can respond to changes in lens temperature because the lens is not perfectly transparent. If the lens is appreciably cooler than the housing, there will be a net transfer of radiant energy from the thermopile to the lens, and the output of the thermopile will not be zero when the pyrometer is sighted on an object whose temperature is the same as that of the thermopile cold junctions.

The heater and resistance thermometer windings referred to above form part of a system devised for maintaining the pyrometer housing and especially the thermopile cold junctions at a constant reference temperature. The resistance thermometer winding is located in a shallow recess in the inner thermopile housing, and a portion of this winding is close to the thermopile cold junctions. This element consists of approximately 1750 ohms of nickel wire at 125°F, and forms the temperature sensitive arm of an a.c. wheatstone bridge. The heater winding is wound as shown in a recess in the outer housing and consists of approximately 800 ohms of suitable heater wire. This winding is in series with the plate circuit of a 2050 thyratron tube. The temperature control circuit (see Fig. 10)[6] consists of an a.c. wheatstone bridge whose output signal is amplified by one stage of voltage gain and applied to the control grid of the 2050 which is a.c. operated. The arrangement is such that a decrease in housing temperature below a preselected value unbalances the bridge in such a direction that the negative bias on the thyratron grid is decreased to the point where the thyratron is allowed to conduct. The resulting thyratron current flows through the

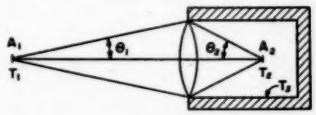

FIG. 8. Simplified radiation pyrometer and source.

[5] T. R. Harrison and W. H. Wannamaker, Rev. Sci. Inst. **12**, 20-32 (1941).

[6] J. C. Mouzon, "A simple temperature controller," Rev. Sci. Inst. **19**, 659 (1948).

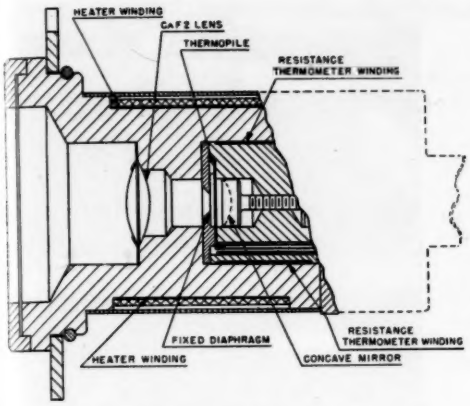

FIG. 9. Low temperature radiation pyrometer.

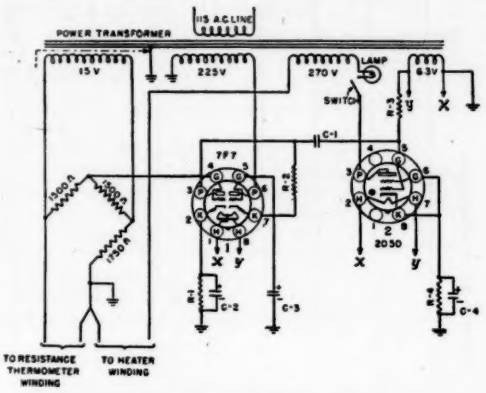

FIG. 10. Circuit diagram of pyrometer housing temperature controller.

heater winding, thus bringing the housing temperature back to the desired control temperature.

The normal control temperature of the housing is chosen to be greater than the highest expected ambient temperature. In this particular case, 125°F has arbitrarily been chosen as the reference temperature. The control sensitivity of the system as used is better than 0.2°F, and, except under extreme conditions of ambient temperature change, the housing temperature is held to ±1°F or better.

Consideration was given during the design of the pyrometer housing to the proper location of the two windings referred to above. It was desirable to locate the resistance thermometer winding as close as practicable to the inner thermopile housing and, at the same time, to maintain a good heat conducting path between it and the heater winding. The first provision eliminates large temperature gradients between the inner thermopile housing and the temperature sensitive winding. The second provision reduces lag between the heater and the heat detector and thereby minimizes the tendency to produce objectionable thermal cycling in the housing.

OPERATION AND PERFORMANCE OF THE PYROMETER

In practice the output of the pyrometer is connected to an electronic self-balancing recording potentiometer which has a minimum full-scale span of 500 microvolts. Figure 11 shows how a standard recorder measuring circuit is modified to obtain a 500-microvolt span. Basically, the recorder has a full scale span of 10 millivolts. The addition of the shunting resistors R_6 and R_7 across the potentiometer circuit permits the comparison of the thermopile e.m.f. with the IR drop in R_6 which is 1/20 of the voltage measured by the standard potentiometer circuit. The measurement

Capacitors				Resistors			
C-1	*C*-2	*C*-3	*C*-4	*R*-1	*R*-2	*R*-3	*R*-4
0.02μf	25μf	8μf	500μf	2500	0.1 meg.	1 meg.	50

made in this manner is truly potentiometric because no current flows in the thermopile circuit at balance. The deflection of the recorder, however, is not a perfectly linear function of the e.m.f. measured, because current is drawn from the bridge circuit. The departure from linearity in this case is less than 0.5 percent of full scale and is of no consequence. Variation in slidewire contact resistance is of little significance in this circuit.

Figure 12 is a response curve showing the output of the pyrometer in microvolts plotted against source temperature. The source of radiation was the blackened side of a metal can containing water. From the slope of the curve the system sensitivity is found to be 4.1 microvolts per degree F at 212°F, and about 2.2 microvolts per degree F at 120°F. This change in sensitivity as a function

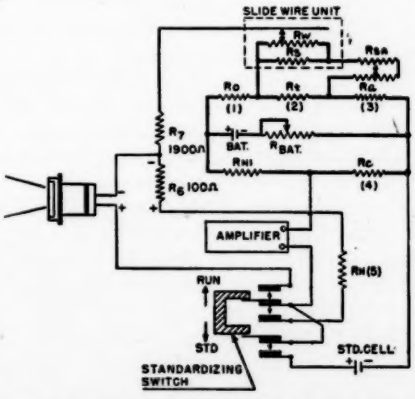

FIG. 11. Potentiometer circuit for low temperature radiation pyrometer (except for R_6 and R_7, components are standard for Brown Electronic Potentiometers).

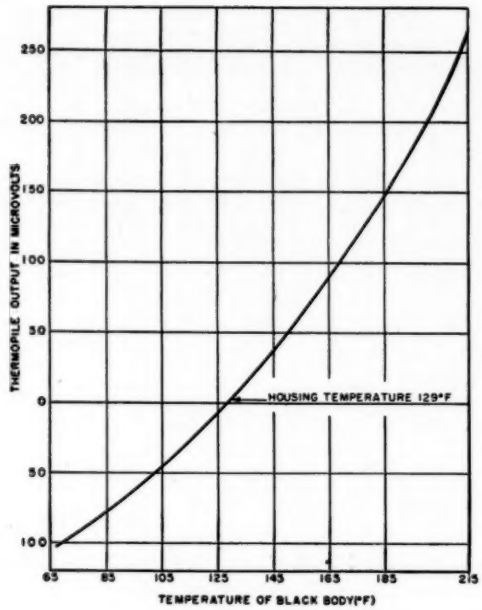

FIG. 12. Blackbody calibration curve.

of temperature is due in part to the change in the transmission of the lens with wave-length.

Figure 13 illustrates an installation where this radiation pyrometer is used to control the temperature of the horizontal rolls in a plastic milling

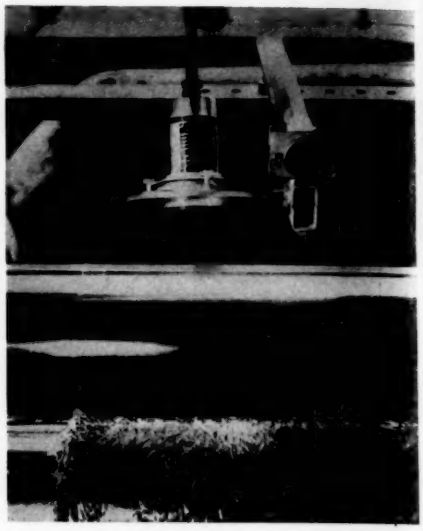

FIG. 13. A field installation of the low temperature radiation pyrometer.

operation. The pyrometer is sighted over the top of the polished roll onto the plastic material.

The authors wish to acknowledge the contributions made by T. R. Harrison and D. G. Tilton to this work, as well as to Mrs. J. R. Clark, who assisted with the measurements and compilation of data.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 39, NUMBER 3 MARCH, 1949

Night Sky Brightness Measurements in Latitudes below 45°

E. O. HULBURT
Naval Research Laboratory, Washington, D. C.
(Received November 12, 1948)

Average values are given of the night sky brightness observed with visual photometers in Brazil, Bikini, Maryland, and New York State, which covered latitudes from 17° south to 45° north. The measurements revealed no important variations with latitude. For clear air, no moon, and no aurorae the brightness decreased on the average from about 225 millimicrolamberts at 15° above the horizon to 130 at the zenith. Previous determinations of the height of the night sky luminosity by means of the observed curves of brightness against zenith angle have lead to conflicting results, values from 60 to more than 500 km being reported. When the determination was repeated with recent data and with the improved corrections of Piotrowski for atmospheric attenuation and scattering such discrepancies between theory and experiment appeared that no certain values of the height were obtained.

THE luminosity of the night sky in places remote from artificial illumination is due to several sources of light, all of which are at a considerable distance. The sources consist of atomic and molecular radiation from regions of the upper atmosphere sometimes termed the polar aurora and the non-polar or permanent aurora, radiations and scattered sunlight from interplanetary space, and radiation and starlight from interstellar space. When one sees or measures the brightness of a place in the sky, he is seeing or measuring the sum of all the radiations which make up the sky luminosity. In a report[1] of 1943, O'Brien and his co-workers of the University of Rochester pointed out that measurements of the visual brightness of the night sky appearing in scientific literature were few and discordant, and described a new series of measurements made at a station in New York State with a visual low brightness photometer devised for the purpose. We have continued the investigation by making similar measurements with photometers patterned after the Rochester instrument at stations in 17° south latitude and in 11° and 39° north latitude. The photometers had a field of view about 11° in diameter with a comparison field illuminated by a radium activated phosphor button, and were calibrated against a non-selectively attenuated source at a color temperature of 2360° Kelvin. From the spectral intensity distribution curve[2] of the non-polar night sky luminosity it appeared that with the photometer and the dark adapted eye one measured chiefly the strong oxygen line 5577 with only minor influence of the sodium yellow line and the molecular lines in the blue and green.

The data are summarized in Fig. 1 in which B millimicrolamberts (mμL) is the brightness of a place in the sky at zenith angle z. (1 mμL$=10^{-9}$ lambert$=2.96\times10^{-7}$ candle per square foot.) The measurements were made when there was no moon, no noticeable aurora and when the sun was more than 18° below the horizon in order to avoid twilight; bright stars were avoided. The four triangles in Fig. 1 were from the data[1] obtained at the Whiteface Mountain Observatory, New York State, latitude 44°30′ north, longitude 73°58′ west and altitude 2500 feet, in clear weather from January 24 to April 23, 1942. The points refer to the zenith and to places in the sky at $z=45°$ to the north, east, south and west. Each value was an average of about 70 readings; individual readings varied not more than about 40 percent from the mean due probably to instrument settings, the variations in haze, and to real variations in B. The fact that the point to the north at $z=45°$ was a little brighter than the other points was probably due to slight aurorae.

The circles in Fig. 1 were data[3] obtained on clear nights at a station near Bocaiuva, Brazil, latitude 17°13′ south, longitude 43°41′ west, and altitude 2200 feet. The data were averages of values taken between 9 and 12 P.M. in directions mostly north around to east during the period from the middle of April to the middle of May, 1947. In the evening the zodiacal light was in the west and the Milky Way ran across the westerly and southerly quadrants; these luminous apparations as well as the brighter stars were avoided in obtaining the values in Fig. 1. Two values for the brightest region of the Milky Way just above the Southern Cross are shown in Fig. 1 by the circles at 263 and 251 mμL. The Magellanic Clouds were visible but not conspicuous in the southwest. In May in the evening the Large Cloud was at an altitude of about 15° and in that position added about 25 mμL to the brightness of the neighboring sky.

[1] OSRD Report *A Series of Measurements of the Brightness of the Night Sky*, and the supplement thereto (Institute of Optics, University of Rochester, 1943).
[2] H. W. Babcock and J. J. Johnson, Astrophys. J. **94**, 271–275 (1941).
[3] R. A. Richardson and E. O. Hulburt, "Brightness of the night sky near Bocaiuva, Brazil." A paper to be published in a technical report of the National Geographic Society.

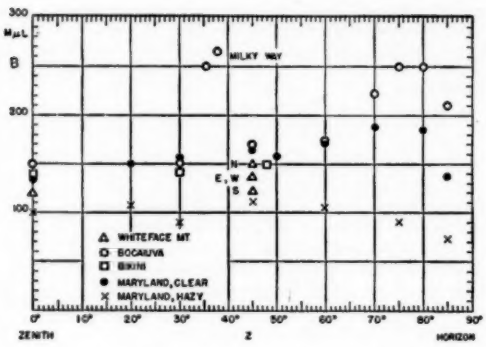

FIG. 1. Night sky brightness at several stations.

A few measurements of B were made in June, 1946, at Bikini atoll, latitude 11°35′ north and 165°30′ east; they are plotted as squares in Fig. 1. Although the lower air was clear, the day horizontal visual range being almost always greater than 20 miles, continuous scattered clouds interferred with the night sky observations. The squares plotted in Fig. 1 were obtained from measurements on clear spaces of sky between the clouds; but for z greater than 45° the clouds, being in echelon, precluded any satisfactory views of the clear sky.

Measurements of B were made at a station in Maryland, a few feet above sea level and at latitude 38°32′ north, longitude 76°10′ west. The station was remote from the illumination of cities and towns. Two sets of the Maryland data are plotted in Fig. 1, the dots being measurements made during two very clear nights in April, 1948, and the crosses being measurements made on typical slightly hazy summer nights in July 1943, when the day horizontal visual range was about 8 miles. It is seen that the effect of haze was to reduce the zenith value of B and to flatten out the maximum in the B,z curve, which is at about $z=75°$ for a clear atmosphere. The effects of haze are in accord with theoretical expectation. The Maryland data were averages of observations made to the southeast, south and southwest. The fact, apparent in Fig. 1, that the maximum in the Maryland B,z curve for a clear atmosphere was below the maximum of the B,z curve for Bocaiuva, was believed to be due mainly to the slightly greater amount of haze in the Maryland air than in the Bocaiuva air. Unquestionably, the night sky toward the horizon was darker at the Maryland station than at Bocaiuva, and as far as memory served the stars in the Maryland sky, although the nights seemed very clear, did not seem as brilliant as those in the Bocaiuva sky.

It was concluded that the data of Fig. 1 offered no indication of consistent variation of B with latitude at the four stations in latitudes from 17° south to 45° north, because the variations were probably due to differing haze conditions. Although many small and complicated intensity variations of the night sky spectrum with time, season, magnetic activity and place in the sky have been observed,[4] these were not revealed by the average visual values of Fig. 1. In connection with the geographical distribution of night sky intensity (probably non-polar aurora) Farnsworth and Elvey[4] concluded that their photographic observations with a glass prism spectrograph indicated no major differences at Bosque Alegre, Argentina, 32° south; Potrerillos, Chile, 27° south; and in the southern New England states; they stated that their data were too meager for much generalization. Rayleigh[4] concluded from his visual observations that the average intensities of the blue, green and red spectral regions at Terling, England, 51° north; Capetown, 34° south; and Camberra, 35° south; were of comparable magnitude, but tended to be highest at Terling and lowest at Capetown. (The foregoing latitudes are geographic, but probably magnetic latitudes may be more significant.)

The effect of stars on night sky brightness observations deserves discussion. From star counts[5] it is known that the total light of all the stars equals that of 1092 stars of visual magnitude $M=1.0$, and that $\frac{4}{5}$ of the light comes from stars of less brightness than about $M=6$ which are not resolved individually by the unaided eye from the surrounding night sky luminosity. If it is assumed that the stars are distributed uniformly over the sky, the equivalent 1092 first magnitude stars produce an illumination at the earth which would be the same as that produced by a uniformly bright sky of $B=32.6$ mμL. The average sky brightness equivalent to all the stars less bright than $M=6$ is $\frac{4}{5}\times 32.6=26$ mμL. This is somewhat less than the observed values of B which, from Fig. 1, are usually greater than 120 mμL. However, the stars are not uniformly distributed over the sky but are concentrated in the galactic region (the Milky Way), and therefore the unresolved stars contribute con-

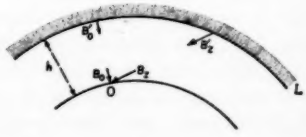

FIG. 2. Geometry of brightness calculation.

[4] The literature of the night sky is voluminous and scattered, see surveys by G. Dejardin, Rev. Mod. Phys. 8, 1–21 (1936); C. T. Elvey, *ibid.* 14, 140–150 (1942); D. R. Bates, M. N. Roy, Astro. Soc. 106, 509–514 (1946); A. Kastler, Ann. de Geophysique, 2, 315–328 (1946).

[5] H. N. Russell, R. S. Dugan, and J. Q. Stewart, *Astronomy* (Ginn and Company, New York, 1926), Vol. 2, p. 625.

siderably to B in some regions of the sky and little in other regions.

When using a visual photometer with no lens, or a weak lens, to measure B the observer can usually avoid regions of very bright stars, but it is rare that he can find regions where there are no stars visually resolved. Thus he sometimes observes B with one or more stars visible in the field of the photometer and makes the instrument setting by putting attention on the general field brightness, thereby consciously discounting to some extent the effect of the visually resolved stars; but the more nearly the stars approach $M=6$, the less of such discounting that is done. With a photoelectric photometer, however, the effect of bright stars may be appreciable, because the instrument records the total illumination which it receives without discrimination as to the source of the illumination. For example, the illumination from a single star of $M=0$, 1 and 2 is equivalent to that from a circular area of sky 5° in diameter of brightness $B=123$, 49 and 20 mμL, respectively.* These values are comparable with the observed values of Fig. 1, which range between 100 and 300 mμL. Therefore, if a photoelectric photometer of 5° diameter field of view is used, data obtained with it should be corrected, or interpreted, for visually resolved stars when compared with data obtained with a visual photometer. However, as yet there are no photoelectric measurements of night sky brightness in visual photometric units.

The height h of the luminous region of the night sky is not known with certainty. The method of determining h depended upon calculation from the observed B,z curve, but the values obtained by various authors[4] ranged from 60 to 500 km and more. Apart from the possibility, or perhaps the probability, that various nocturnal radiations originate at different levels, there were two factors which contributed to the conflicting results, the variability of the observed B,z curve and the lack of an adequate theoretical treatment of the attenuation and scattering of the night sky light along the observer's line of sight. Early writers assumed that the light was attenuated exponentially, as the light from a point source. Later[6] it was pointed out that this was incorrect because one is dealing with a luminous area and an atmosphere which attenuates mainly by scattering with little true absorption. Recently, Piotrowski[7] published a more complete calculation of the scattering effects

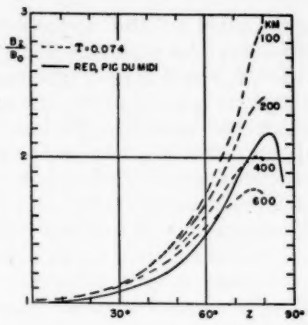

FIG. 3. Comparison of theory with data at Pic du Midi for red light.

using the Eddington approximation in the transfer equations. It was hoped that a discussion of the B,z curves in the light of Piotrowski's equations would lead to more definite values of h, but as shown in the following paragraphs this hope has not been realized.

Referring to Fig. 2 it is assumed that the luminous layer L is a single layer at a considerable height h and is above the lower part of the atmosphere where the scattering and attenaution mainly occur. L is assumed to be uniform and concentric with the surface of the earth; if L is a thick layer, h is an effective height. Let O be the position of the observer and let B_0 and B_z be the brightness of his sky in the zenith and in a place of zenith angle z, respectively. Let the corresponding brightness of L be B_0' and Bz'. Then B_0 is the sum of the light B_0' after being attentuated by the atmosphere plus the light which the atmosphere scatters; similarly for B_z.

If the brightness of the luminous layer is proportional to its thickness in the line of sight, one has[8]

$$B_z'=B_0'[1-(\sin z)^2/(1+h/R)^2]^{-\frac{1}{2}}, \qquad (1)$$

where R is the radius of the earth. From (1) and Piotrowski's equations (3) and (5) of reference 7, one obtains

$$B_z/B_0'=[1-(\sin z)^2/(1+h/R)^2]^{-\frac{1}{2}}e^{-\tau\sec z}$$
$$+\frac{2H}{4+s\tau}[2+3\cos z-(2+3\cos z+s\tau)e^{-\tau\sec z}], \qquad (2)$$

where

$$H=(1+h/R)\{1+h/R-[2h/R+(h/R)^2]^{\frac{1}{2}}\}. \qquad (3)$$

τ is the attenuation coefficient of the air per atmosphere. Thus $e^{-\tau}$ is the fraction of collimated light from a point source in the zenith outside the atmosphere which reaches the observer; τ is assumed to be due entirely to scattering. s is a function of the

* The calculation assumed that for the sun M was -26.72, that the solar illumination just outside the terrestrial atmosphere was 13,600 foot candles, and that the visible spectral energy distribution of the star was the same as that of the sun.

[6] D. Barbier, Ann. de geophysique 1, 144–156 (1944); E. O. Hulburt, report at the Gassiot Committee conference at London, July, 1947.

[7] S. L. Piotrowski, Astrophys. J. 106, 466–471 (1947).

[8] P. J. van Rhijn, Pub. Astr. Lab. Groningen, No. 21 (1921).

scattering with angle of the atmosphere. From his unpublished data for clear air Piotrowski obtained $s=2.4$ for visual rays. Other data[9] on atmospheric angular scattering do not agree very well with $s=2.4$, but probably the discrepancy is not serious for τ less than, say, 0.3. Since earth curvature and atmospheric refraction were partially neglected in the derivation of (2), as evidenced by the appearance of sec$z$ in the exponents, (2) becomes inaccurate for $z>80°$. Piotrowski showed that his equations gave fair agreement with certain of Abbot's[10] measurements of daylight sky intensities. The first term of the right side of (2) is the contribution to B_z from the night sky emissions in the line of sight which reach the observer after attenuation by the atmosphere, and the second term is the contribution from the emissions over the entire hemisphere which are scattered into the line of sight. As might be expected, the relative importance of the second term increases with τ and z; for $\tau=0$ the second term is zero and (2) reduces to (1) as it should. For $z=0$, $B_z \doteq B_0$ and (2) reduces to

$$B_0/B_0' = e^{-\tau} + \frac{2H}{4+s\tau}[5-(5+s\tau)e^{-\tau}]. \qquad (4)$$

The ratio B_z/B_0 is found by dividing (2) by (4); it is this ratio that is to be compared with the observations.

Now Piotrowski's derivation omitted the effects of light reflected upward from the surface of the earth and of light from other sources than the upper air luminosity, as the stars. These two streams of light are scattered in the lower atmosphere and increase the second term of the right hand side of (2) by an amount which increased with τ. An approximate estimate indicated that the increase became appreciable (greater than a few percent) when $\tau>0.3$, but since such a case, as shown below, was not useful for the determination of h, the exact evaluation of the increase was not attempted.

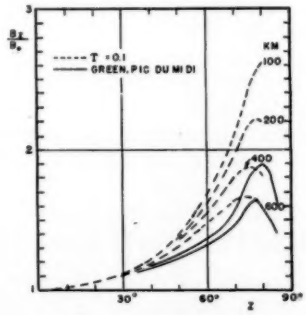

FIG. 4. Comparison of theory with data at Pic du Midi for green light.

The theory was first applied to some careful measurements made by P. Abadie, A. Vassy, and E. Vassy[11] at the Pic du Midi Observatory (north latitude 42°56′) with a photo-multiplier photometer of field of view 6° in diameter with a red and a green filter. Their average normal curve of B_z/B_0 for the red filter is shown in Fig. 3. It was an average of 10 curves which varied from the average by about ±10 percent. Abnormal curves also were recorded which ascended to values of B_z/B_0 between 4 and 5 at $z=80°$. It was believed that the abnormal curves indicated the presence of polar aurorae. A value $\tau=0.074$ at the observatory for red light was stated, based on calculation from earlier observations. The theoretical curves of B_z/B_0 from (2) and (4) with $\tau=0.074$ and $h=100$, 200, 400 and 600 km were plotted in Fig. 3. The observed curve did not agree with these; although it fell among the curves for $h>300$ km, the deviations were such as to render uncertain any inferences about the exact value of h.

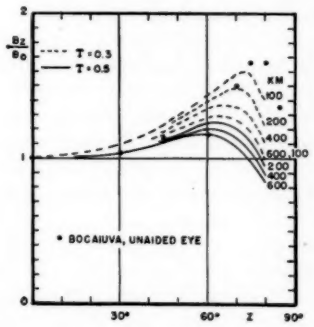

FIG. 5. Comparison of theory with data at Bocaiuva.

The Pic du Midi data[11] with the green filter are shown in the two curves of Fig. 4; it was stated that these were two examples from a large number of normal curves. Many abnormal curves were also observed which were attributed to polar aurora conditions; in some cases the aurorae were clearly visible and in other cases they were not. The green filter data were more erratic than the red; sometimes the abnormalities of the green and red coincided, and sometimes they did not, indicating that the phenomena were complex. A value 0.10 was given for τ for green light based on calculation from earlier observations. The curves for B_z/B_0 calculated from (2) and (4) for $\tau=0.1$ are in Fig. 4, and when compared with the two ob-

[9] E. O. Hulburt, J. Opt. Soc. Am. **31**, 467–476 (1941); A. Bullrich, Optik **2**, 301–325 (1947).
[10] C. G. Abbot, Astrophys. J. **28**, 129–136 (1914).
[11] P. Abadie, A. Vassy, and E. Vassy, Ann. de Geophysique **1**, 189–224 (1944).

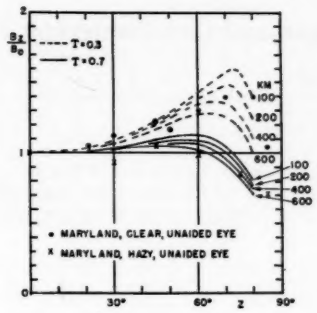

FIG. 6. Comparison of theory with data at Maryland.

served curves poor agreement was found just as in the case of Fig. 3. P. Abadie, A. Vassy, and E. Vassy[11] attempted to improve the agreement by assuming two luminous regions one at 70 km and one at 1000 km; their theoretical curves were calculated from the exponential attenuation theory, which is an imperfect theory; actually they used the first term of the right-hand side of (2).

The Bocaiuva data of Fig. 1 converted into B_z/B_0 were plotted in the dots of Fig. 5. τ was determined in the day from many measurements of the solar illumination and at night from a few measurements of the lunar illumination for several zenith angles of the sun and moon. Values of τ from 0.3 to 0.5 were obtained. The sets of B_z/B_0 curves calculated from (2) and (4) for these values of τ were plotted in Fig. 5. It is seen that none of the curves agreed well with the observed dots. It is also seen that for τ greater than, say, 0.3 the curves for various values of h were compressed together. Therefore, to arrive at a definitive value of h would have required a precision in the B_z/B_0 values which was not approached.

Finally, the Maryland values of Fig. 1 reduced to B_z/B_0 were plotted in Fig. 6. τ was not measured but the two sets of B_z/B_0 curves for $\tau=0.3$ and 0.7 were plotted in Fig. 6. Since these curves were not unlike the trends of the observed points, it appeared that the theory gave a fair account of the conditions in a hazy atmosphere, but since the curves were closely spaced together no estimate of h was derived from them.

It was concluded that a value of h could not be derived with certainty from the present data, because the agreement between theory and experiment was not completely satisfactory. Further progress might be made by carrying out measurements of B_z/B_0 at high altitudes, as from airplanes, balloons or rockets, in order to reduce as much as possible the corrections for atmospheric scattering. This might be better than to attempt to improve the corrections by a more refined calculation, and by more complete and accurate observations of B_z and τ at low altitudes. It might be advantageous to work at high altitudes in low latitudes to minimize the complicating effects of polar aurorae, because the relations between the polar aurorae and non-polar or permanent aurora are not clear. Instrumental methods with photocells for measuring light intensities are more accurate and rapid than visual methods. The different portions of the spectrum should be examined separately, for it is almost certain that the various radiations originate in different levels. It may turn out that the night luminosity of the upper atmosphere is too variable or too non-uniformly distributed to permit a precise determination of the altitude from the brightness zenith angle curve.

The Infra-Red Emission Spectra of the Oxy-Hydrogen and Oxy-Deuterium Flames*

S. SILVERMAN AND R. C. HERMAN
Applied Physics Laboratory, The Johns Hopkins University, Silver Spring, Maryland
(Received November 10, 1948)

The infra-red emission spectra of the oxy-hydrogen and oxy-deuterium flames and of hot water vapor have been studied in the LiF, NaCl, and KBr regions. The principal features of the spectrum may be accounted for by emission of H_2O molecules; no unequivocal evidence for radiation from OH radicals was found with the resolving power available.

IN the past there have been numerous investigations of the infra-red emission from flames of various types. Recent improvements in infra-red techniques have made it seem worth while to repeat some of this earlier work. In particular the oxy-hydrogen flame is of interest and, in fact, has been recently studied by Plyler and Humphreys.[1] The purpose of this note is to report similar investigations made by us.

The infra-red emission spectra of the oxy-hydrogen and oxy-deuterium flames have been investigated in the prism region with an automatic recording Perkin-Elmer spectrometer. The oxy-hydrogen flame spectrum was studied in the region 500 to 7500 cm^{-1} using LiF, NaCl, and KBr optics, whereas the oxy-deuterium flame was studied in the LiF region only, for reasons of economy. All the measurements were carried out with an a.c. amplification system, except for those in the LiF region where d.c. amplification was employed to obtain higher sensitivity. The gases were premixed and burned in an acetylene type torch. A number 2 or 3 orifice was used with hydrogen in order to obtain a large, hot flame. On the other hand, the oxy-deuterium mixture was burned in a number 00 orifice to conserve deuterium. The flames were burned in a water-cooled flue to improve stability, and the amount of radiation entering the spectrometer was increased by means of a concave mirror placed behind the flame. Atmospheric absorption was minimized by continuously flushing the spectrometer with dry nitrogen.

The emission spectra of the oxy-hydrogen flame in the LiF, NaCl, and KBr regions are shown in Figs. 1a, 2, and 3, respectively. These as well as the other spectra shown are unretouched tracings of typical automatic recordings. Figure 4 shows the emission from the oxy-deuterium flame in the LiF region. The effective slit widths employed are indicated in the figures. It is to be noted that spectra in the 3μ region exhibit prism absorption, principally at 2.813, 2.790, 2.755, and 2.670μ, which is caused by a hydride impurity.[2]

In Table I are collected the positions of the emission bands observed in both the oxy-hydrogen and oxy-deuterium flames. The bands from the oxy-hydrogen flame correspond, in the main, with the well known H_2O absorption bands. The emission bands from the oxy-deuterium flame seem also to be attributable to hot D_2O. Most of the D_2O emission bands given in Table I show roughly the correct isotope shift from the H_2O bands. The emission bands do not correspond exactly to the absorption bands because they are temperature broadened and

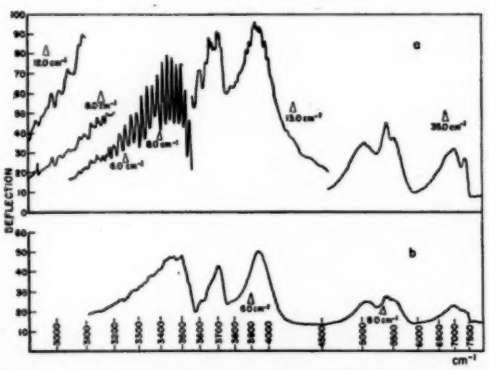

FIG. 1. (a) The emission spectrum of the oxy-hydrogen flame in the region 2950–7500 cm^{-1}. Slit widths are indicated by the numbers next to the Δ's. (b) The emission spectrum of a jet of hot water vapor in the LiF region. (The temperature is roughly 1300°K.) These curves both show the absorption in the region of 3600 cm^{-1} which is characteristic of this LiF prism. Slit widths are indicated by the numbers next to the Δ's.

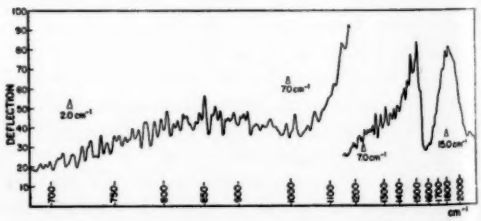

FIG. 2. The emission spectrum of the oxy-hydrogen flame in the region 700–2000 cm^{-1} obtained with a NaCl prism. Slit widths are indicated by the numbers next to the Δ's.

* The work described in this paper was supported by the Bureau of Ordnance, U. S. Navy, under Contract NOrd-7386.

[1] E. K. Plyler and C. J. Humphreys, J. Research Nat. Bur. of Stand. **40**, 449 (1948).

[2] N. Wright, Rev. Sci. Inst. **15**, 22 (1944).

TABLE I. Observed emission bands.

H_2O	D_2O
7380 cm^{-1}	5530 cm^{-1}
7040	5350 5260
5525	5175
5385	4110
5080	4005
3890	3760
3705	2885
	2775

TABLE II. Observed oxy-hydrogen and oxy-deuterium flame lines.

Oxy-hydrogen	Oxy-hydrogen	Oxy-hydrogen	Oxy-hydrogen	Oxy-deuterium
3545 cm^{-1}	2940 cm^{-1}	956 cm^{-1}	722 cm^{-1}	2725 cm^{-1}
3527	2905	942	716	2690
3503	2885	929	708	2675
3482	2862	915	699	2665
3460	2833	907	697	2650
3437	2820	904	693	2635
3415	1830	893	685	2625
3393	1507	885	677	2612
3365	1460	880	673	2605
3345	1440	872	661	2593
3317	1420	867	650	2582
3295	1377	861	642	2572
3277	1356	852	637	2558
3257	1338	846	629	2545
3230	1314	839	619	2537
3205	1288	832	611	2525
3185	1265	828	606	2515
3178	1255	824	595	2503
3165	1243	817	592	2488
3147	1234	812	587	2480
3142	1220	804	584	2465
3128	1208	798	578	2452
3123	1200	791	567	2443
3103	1188	782	560	2430
3097	1179	774	553	2417
3077	1158	768	542	2407
3057	1147	761	537	2395
3040	1120	755	530	2355
3030	1065	750	523	
3015	1044	744	517	
2990	1013	739	511	
2970	991	734	505	
2955	967	729	497	

there is undoubtedly some emission involving upper states. We have also obtained the emission from

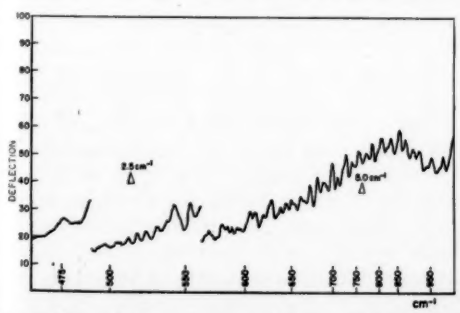

FIG. 3. The emission spectrum of the oxy-hydrogen flame in the region 465–975 cm^{-1} obtained with a KBr prism. Slit widths are indicated by the numbers next to the Δ's.

TABLE III. Observed H_2O absorption lines.

Sleator and Plyler*,** vibration-rotation lines	Weber and Randall** vibration-rotation and pure rotation lines	Ginsburg** pure-rotation lines
3503 cm^{-1}	960 cm^{-1}	618 cm^{-1}
3494	904	611
3484	891	605
3464	883	600
3449	866	592
3444	863	585
3433	858	581
3422	850	576
3411	835	564
3394	824	561
3367	810	550
3357	806	540
3338	801	536
3314	795	534
3299	790	526
3292	780	520
3289	774	515
3277	766	502
3255	756	498
3245	751	492
3237	741	
3231	728	
3197	711	
3186	703	
3179	695	
3128	692	
3123	688	
3115	678	
3108	660	
3102	652	
3097	650	
3079	643	
3068	640	
3065	635	
3057	628	
3050	625	
3035		
3033		
3018		
3012		
1830		
1507		
1472		
1457		
1465		
1437		

* This table contains the data of Sleator and Plyler rather than the data from the more recent work of H. H. Nielsen, Phys. Rev. **59**, 565 (1941) and **62**, 422, (1942), since our data should not be compared with spectra taken at such high resolution.
** See reference 3.

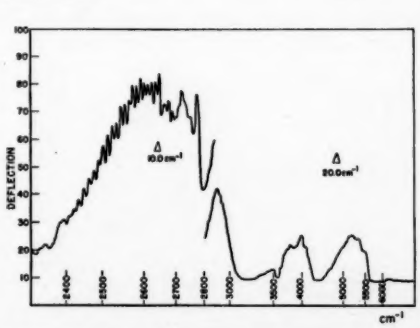

FIG. 4. The emission spectrum of the oxy-deuterium flame in the region 2350–6500 cm^{-1} obtained with a LiF prism. Slit widths are indicated by the numbers next to the Δ's.

TABLE IV.* Vibration-rotation lines of OH calculated.

2–1	1–0
2993 cm^{-1}	3095 cm^{-1}
2946	3046
2898	2995
2849	2944
	2892
	2839

* All the positions listed are averages of a P_1 and a P_2 line.

TABLE V. Rotation lines of OH calculated.

989 cm^{-1}	765 cm^{-1}
970	738
950	711
930	683
908	654
886	624
867	594
844	563
816	531
791	498

hot H_2O gas in the LiF region. This spectrum is shown in Fig. 1b for comparison.

It would be highly desirable to study these spectra under the high resolution afforded by a grating instrument. Nevertheless, we have measured and recorded in Table II all the vibration-rotation and pure rotation lines observed in both the oxy-hydrogen and oxy-deuterium flames. It should be pointed out that the observed frequencies are accurate to about 4 or 5 cm^{-1}, and represent averages of several spectral runs. For comparison we also list in Tables III, IV, and V some of the principal H_2O absorption lines[3] as well as the OH lines computed by Dieke and Crosswhite.[4]

An inspection of the tables shows within the experimental error frequent agreement between the observed emission lines and the observed absorption lines of the H_2O molecule. It is to be noted that practically all of the lines that might be attributed to the OH radical could as easily be attributed to the H_2O molecules. This is true of the rotation lines in the 1−0 and less intense 2−1 vibration-rotation bands of the OH radical as well as the pure rotation lines of this radical. It is our opinion that with the present resolution, our spectra, although somewhat better resolved than hitherto, do not give unequivocal evidence of the existence of OH radiation in the region of the spectrum studied. In conclusion, the flame spectra examined, under prism resolution, show the general character to be expected from emission by water molecules. It would appear that positive identification of the OH spectrum in the infra-red region of the spectrum will require higher resolving power.

The authors wish to thank Dr. N. Ginsburg, Syracuse University, for discussing with them the results of his calculations on the pure rotation spectrum of the H_2O molecule prior to publication. In conclusion we express our appreciation to the U. S. Atomic Energy Commission for furnishing us with 99.3 percent pure deuterium gas.

[3] These data were obtained from: E. K. Plyler and W. W. Sleator, Phys. Rev. **37**, 1493 (1931); L. R. Weber and H. M. Randall, Phys. Rev. **40**, 835 (1932); N. Ginsburg, Phys. Rev. **74**, 1052 (1948).

[4] G. H. Dieke and H. M. Crosswhite, *The Ultraviolet Bands of OH*, Johns Hopkins University, CF-790.

Cadmium Spectra

A. G. SHENSTONE AND J. T. PITTENGER*
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received November 8, 1948)

The hollow cathode spectrum of cadmium has been observed over the wave-length range from 600A to 6500A. In Cd I the ^3D series has been resolved and extended. In Cd II the work of von Salis and Takahashi has been corrected and considerably extended. The ^2P and ^2F series are very perturbed by the overlapping $4d^95s5p$ configuration which is nearly complete. No series terms have been found in the $4d^95sns$ series. Cd III is excited very strongly in the discharge and the analysis is now much extended. The odd terms are from $4d^95p$ and they combine with $4d^{10}$, $4d^95s$, $4d^96s$, $4d^95d$, and $4d^95s^2$. An ionization potential of 37.6 volts can be calculated from the $4d^9ns$ series.

THE term table of the second spectrum of cadmium which appears in Bacher and Goudsmit's "Atomic Energy States" is based on papers by G. von Salis[1] and Y. Takahashi[2] which appeared in 1925 and 1929, respectively. The first of these papers was confined to an analysis of that part of the spectrum which does not require observation below 2000A and dealt with the more or less regular doublet spectrum that is to be expected from the cadmium ion. The later paper extended the observations to the vacuum region and was concerned with the more complicated quartet-doublet structures such as occur in the similar spectra Cu I and Ag I. Examination of Takahashi's lists of observed lines indicated that his analysis could not be wholly correct since they included a number of lines definitely known to belong to Cd III. We therefore undertook new observations of a more accurate and comprehensive nature. Both von Salis and Takahashi had used a hollow cathode tube, a source which is normally expected to produce lines whose excitation is limited by the noble gas present. This limitation confines the excitation to a maximum of 198,308 cm^{-1} above the ground state of the atom when helium is used and to smaller values for the other noble gases. In cadmium, with helium as the gas, this limit lies 10,600 cm^{-1} below the second ionization and should therefore lead to an emitted spectrum including no Cd III lines at all. In point of fact, this expectation is far from being realized since lines of Cd III are emitted which require excitation of 213,108 cm^{-1} above the second ionization point, or a total of 72,535+136,375+213,108=422,018 cm=52.1 volts. This extreme excitation occurs with either helium or neon in the tube and is probably due to the high vapor pressure of the metal which exists even in a water-cooled cathode and consequent electron excitation. There is a small possibility that the excitation is due to collisions of doubly ionized gas atoms with cadmium atoms but this would yield far too much energy.

The type of tube we used for our observations had a water-cooled hollow cathode, the cadmium metal being contained in a thin graphite or iron cylinder. The power was supplied by a rectifier fed from a constant current circuit and the tube was operated in the region of 400 to 500 volts and 0.4 to 0.9 ampere. Plates were taken on a 21-ft. 30,000-line per inch grating down to 2200A and beyond that point on a 2-meter 30,000-line per inch grating. The exposure times ranged from a few minutes to two hours. The standards used were iron arc lines in the air region and copper lines in the vacuum region. In the longer wave-lengths, disagreements between observed and calculated wave numbers greater than ± 0.1 cm^{-1} rarely occur, but in the Schumann region occasional errors of as much as ± 2 cm^{-1} are present.

In addition to the spectra of the hollow cathode, some exposures to sparks in nitrogen were made on the vacuum spectrograph. These and the observations of L. and E. Bloch[3] and F. Esclangon[4] of the spectra of the electrodeless discharge were very valuable in the differentiation of the degrees of ionization. In the spark, transitions from middle to low levels of Cd II appear sharp and strong, transitions from high to middle levels are weak if present at all. On the other hand, transitions in

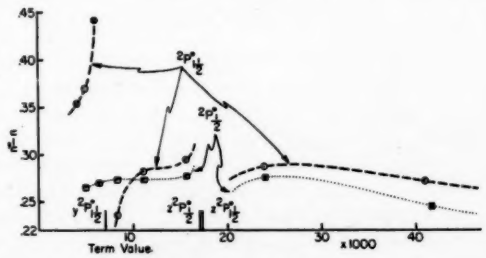

FIG. 1. Perturbed $^2P^\circ$ series.

* A considerable part of the material in this paper is from the senior thesis of J. T. P. written in 1940.

[1] G. von Salis, Ann. d. Physik **76**, 145 (1925).

[2] Y. Takahashi, Ann. d. Physik **3**, 27 (1929).

[3] L. and E. Bloch, Ann. de Physique **5**, 325 (1936).

[4] F. Esclangon, J. de phys. et rad. **7**, 52 (1926).

TABLE I. ^3D terms of cadmium I.

N	$^3D_1$	Δ	$^3D_2$	Δ	$^3D_3$
7	4548.7	2.8	4545.9	4.6	4541.3
8	3138.3	2.3	3136.0	2.7	3133.3
9	2294.6	1.3	2293.3	1.6	2291.7
10	1750.5	0.7	1749.8	1.1	1748.7
11			1379.2	1.0	1378.2
12			1114.7	0.9	1113.8

Cd III from high to middle levels are strong and diffuse. An intermediate diffuseness is shown by a considerable number of unidentified lines which are probably the transitions in Cd III from $4d^85s5p$ to $4d^85s^2$.

CADMIUM I

Most of the known lines of the cadmium arc spectrum were observed and a number of triplet terms were resolved for the first time. The term values based on the new measurements are given in Table I.

CADMIUM II

The basic analysis of this spectrum is that of von Salis, and it is mainly confirmed by the new observations. The ^2S and ^2D series have been extended and afford the means for an accurate calculation of the limit of all the ordinary series. As remarked by von Salis, the $^2P°$ and $^2F°$ series are both very irregular. The new observations and the more complete analysis, have served to correct his assignments and to extend the series which now show the characteristics of doubly perturbed series. They are shown in Figs. 1 and 2 in which n^*-n is plotted against term value. It is to be noted that perturbations of comparable magnitude occur in both series whereas in the similar spectrum Cu I[5] the $^2P°$ series is radically upset and the $^2F°$ series is practically normal. The perturbing terms are the two $^2F°$ terms and the two $^2P°$ terms from $4d^95s5p$. The lower of these terms in each case cause a small irregularity in the run of the series whereas the upper ones have a very marked effect. This is true in Cu I also and perhaps shows that the higher terms are more nearly pure doublets than the lower ones which are part of a doublet-quartet group.

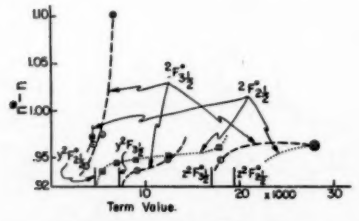

FIG. 2. Perturbed $^2F°$ series.

The additions to von Salis's analysis which were made by Takahashi were as follows. He added the very important $4d^95s^2\ ^2D$ term and he found a number of odd levels, "A" levels, which combine with it and with other even levels. Most of these are real though incorrectly identified. On the other hand he lists a group of high even levels, "B" levels, all of which must be rejected as unreal. He expresses doubt of their reality.

Takahashi's "A" levels should include the many levels of $4d^95s5p$ comprising $^4P°\ ^4D°\ ^4F°$ and two sets of $^2P°\ ^2D°\ ^2F°$. These levels are mixed in with the $^2P°$ and $^2F°$ series and the mutual perturbations are evident in the irregularities in the series and in the intensities of the combinations. It is possible, however, to name unambiguously most of the levels of the group. The only exceptions are as follows. The $^4F_{4\frac{1}{2}}°$ cannot be found because no level is known with which it can combine. The $^4D_{3\frac{1}{2}}°$ is doubtful because it is based on one combination only, but that one line is the only remaining line in the correct region of the spectrum. In addition a comparison with the same structure in Ag I[6] makes the identification very probable. The upper set of doublets lacks $^2P_{\frac{1}{2}}°$ and both components of $^2D°$.

Takahashi's rejected "B" levels all lie above the limit of the ordinary series and extend into the region where one might expect the structure $4d^95s6s$. The position of this important group can be predicted with some certainty from the fact that it is really in series with $4d^95s^2$. The calculation places it at above 22,400 cm^{-1} above $4d^{10}\ ^1S_0$, and no other even levels will lie below it. This eliminates most of the "B" levels. They are all in fact, based on far too large tolerances of error. A very careful search was made for the levels of $4d^95s6s$, which should produce strong multiplets in combination with $4d^95s5p$. As in Cu I there may be considerable auto-ionization to reduce the intensity in a low pressure source such as was used, but this does not apply to $^4D_{3\frac{1}{2}}$ and $^4D_{\frac{1}{2}}$ which should find no continua into which to ionize. There are, however, no unidentified lines where the chief multiplets should fall. It appears that the excitation in the hollow cathode does not extend into the region above the series limit $4d^{10}$ of Cd II. There does exist much higher excitation but it involves the removal of another electron and the spectrum is then Cd III, not Cd II.

The intense excitation in a hollow cathode discharge using helium usually extends to the ionization potential of helium above the lowest state of the atom. This point is 125,774 cm^{-1} in the Cd II

[5] A. G. Shenstone, Phil. Trans. Roy. Soc. **A241**, 297 (1948).

[6] A. G. Shenstone, Phys. Rev. **57**, 894 (1940).

TABLE II. Terms of cadmium II.

1	2	3	4	5
$4d^{10}5s$	$5s\,^2S_{\frac{1}{2}}$	00.00	1.7941	s
$4d^{10}5p$	$5p\,^2P^\circ_{\frac{1}{2}}$	*44136.08*	2.1815	s
$4d^{10}5p$	$5p\,^2P^\circ_{1\frac{1}{2}}$	*46618.55*	2.2114	s
$4d^95s^2$	$a\,^2D_{2\frac{1}{2}}$	69258.91		t
$4d^95s^2$	$a\,^2D_{1\frac{1}{2}}$	74893.66		t
$4d^{10}6s$	$6s\,^2S_{\frac{1}{2}}$	82990.66	2.8675	s
$4d^{10}5d$	$5d\,^2D_{1\frac{1}{2}}$	89689.25	3.0663	s
$4d^{10}5d$	$5d\,^2D_{2\frac{1}{2}}$	89843.78	3.0714	s
$4d^{10}6p$	$6p\,^2P^\circ_{\frac{1}{2}}$	*94710.41*	3.2458	s
$4d^{10}6p$	$6p\,^2P^\circ_{1\frac{1}{2}}$	*95383.63*	3.2723	s
$4d^95s5p$	$z\,^4P^\circ_{2\frac{1}{2}}$	*106189.20*		
$4d^{10}7s$	$7s\,^2S_{\frac{1}{2}}$	107300.88	3.8856	t
$4d^{10}4f$	$4f\,^2F^\circ_{2\frac{1}{2}}$	*108419.47*	3.9623	s
$4d^{10}4f$	$4f\,^2F^\circ_{3\frac{1}{2}}$	*108432.43*	3.9635	s
$4d^95s5p$	$z\,^4P^\circ_{1\frac{1}{2}}$	*109440.86*		t
$4d^95s5p$	$z\,^4F^\circ_{3\frac{1}{2}}$	*109736.82*		
$4d^95s5p$	$z\,^4F^\circ_{2\frac{1}{2}}$	*110168.11*		t
$4d^{10}6d$	$6d\,^2D_{1\frac{1}{2}}$	110174.10	4.0931	s
$4d^{10}6d$	$6d\,^2D_{2\frac{1}{2}}$	110247.60	4.0988	s
$4d^95s5p$	$z\,^4P^\circ_{\frac{1}{2}}$	*112196.50*		
$4d^{10}7p$	$7p\,^2P^\circ_{\frac{1}{2}}$	*112361.06*	4.2751	t
$4d^{10}7p$	$7p\,^2P^\circ_{1\frac{1}{2}}$	*112490.39*	4.2870	t
$4d^95s5p$	$z\,^4F^\circ_{1\frac{1}{2}}$	*112785.19*		
$4d^95s5p$	$z\,^4D^\circ_{3\frac{1}{2}}$	*114793.91?*		
$4d^95s5p$	$z\,^4D^\circ_{2\frac{1}{2}}$	*114952.30*		t
$4d^95s5p$	$z\,^4D^\circ_{1\frac{1}{2}}$	*116225.71*		t
$4d^95s5p$	$z\,^2F^\circ_{2\frac{1}{2}}$	*116953.70*		t
$4d^95s5p$	$z\,^4D^\circ_{\frac{1}{2}}$	*117988.79*		t
$4d^{10}8s$	$8s\,^2S_{\frac{1}{2}}$	118040.65	4.8930	s
$4d^{10}5f$	$5f\,^2F^\circ_{3\frac{1}{2}}$	*118444.36*	4.9478	
$4d^{10}5f$	$5f\,^2F^\circ_{2\frac{1}{2}}$	*118545.77*	4.9618	
$4d^95s5p$	$z\,^2P^\circ_{1\frac{1}{2}}$	*119056.11*		t
$4d^95s5p$	$z\,^2P^\circ_{\frac{1}{2}}$	*119293.99*		
$4d^95s5p$	$z\,^2F^\circ_{3\frac{1}{2}}$	*119373.19*		
$4d^{10}7d$	$7d\,^2D_{1\frac{1}{2}}$	119522.72	5.1036	s
$4d^{10}7d$	$7d\,^2D_{2\frac{1}{2}}$	119561.99	5.1090	s
$4d^95s5p$	$z\,^2D^\circ_{1\frac{1}{2}}$	*120135.22*		t
$4d^{10}8p$	$8p\,^2P^\circ_{\frac{1}{2}}$	*120618.46*	5.2782	
$4d^{10}8p$	$8p\,^2P^\circ_{1\frac{1}{2}}$	*120711.18*	5.2937	s
$4d^95s5p$	$z\,^2D^\circ_{2\frac{1}{2}}$	*121281.64*		t
$4d^{10}9s$	$9s\,^2S_{\frac{1}{2}}$	123751.48	5.8969	s
$4d^{10}6f$	$6f\,^2F^\circ_{3\frac{1}{2}}$	*123972.10*	5.9490	s
$4d^{10}6f$	$6f\,^2F^\circ_{2\frac{1}{2}}$	*123988.39*	5.9530	s
$4d^{10}6g$	$6g\,^2G$	124151.26	5.9925	s
$4d^{10}8d$	$8d\,^2D_{1\frac{1}{2}}$	124613.30	6.1091	s
$4d^{10}8d$	$8d\,^2D_{2\frac{1}{2}}$	124636.50	6.1151	s
$4d^{10}9p$	$9p\,^2P^\circ_{\frac{1}{2}}$	*125223.06*	6.2739	t
$4d^{10}9p$	$9p\,^2P^\circ_{1\frac{1}{2}}$	*125254.49*	6.2827	t
$4d^{10}10s$	$10s\,^2S_{\frac{1}{2}}$	127152.89	6.8991	s
$4d^{10}7f$	$7f\,^2F^\circ_{3\frac{1}{2}}$	*127249.09*	6.9368	s
$4d^{10}7f$	$7f\,^2F^\circ_{2\frac{1}{2}}$	*127283.04*	6.9483	s
$4d^{10}7g$	$7g\,^2G$	127396.62	6.9922	s
$4d^{10}9d$	$9d\,^2D_{1\frac{1}{2}}$	127697.43	7.1124	s
$4d^{10}9d$	$9d\,^2D_{2\frac{1}{2}}$	127712.28	7.1185	s
$4d^{10}10p$	$10p\,^2P^\circ_{1\frac{1}{2}}$	*127996.73*	7.2383	t
$4d^{10}10p$	$10p\,^2P^\circ_{\frac{1}{2}}$	*128076.3*	7.2730	
$4d^95s5p$	$y\,^2F^\circ_{2\frac{1}{2}}$	*129189.5*		
$4d^{10}11s$	$11s\,^2S_{\frac{1}{2}}$	129343.05	7.9009	s
$4d^95s5p$	$y\,^2P^\circ_{1\frac{1}{2}}$	*129344.86*		
$4d^{10}8f$	$8f\,^2F^\circ_{2\frac{1}{2}}$	*129419.38*	7.9441	t
$4d^{10}8g$	$8g\,^2G$	129502.43	7.9921	s
$4d^{10}8f$	$8f\,^2F^\circ_{3\frac{1}{2}}$	*129686.60*	8.1014	
$4d^{10}10d$	$10d\,^2D_{1\frac{1}{2}}$	129708.50	8.1146	s
$4d^{10}10d$	$10d\,^2D_{2\frac{1}{2}}$	129718.66	8.1207	s
$4d^{10}11p$	$11p\,^2P^\circ_{\frac{1}{2}}$	*129956.3*	8.2697	
$4d^{10}11p$	$11p\,^2P^\circ_{1\frac{1}{2}}$	*130216.6*	8.4426	
$4d^{10}12s$	$12s\,^2S_{\frac{1}{2}}$	130835.86	8.9024	s
$4d^{10}9f$	$9f\,^2F^\circ_{2\frac{1}{2}}$	*130878.61*	8.9369	
$4d^{10}9f$	$9f\,^2F^\circ_{3\frac{1}{2}}$	*130925.82*	8.9753	
$4d^{10}9g$	$9g\,^2G$	130946.09	8.9922	
$4d^{10}11d$	$11d\,^2D_{1\frac{1}{2}}$	131092.90	9.1162	t
$4d^{10}11d$	$11d\,^2D_{2\frac{1}{2}}$	131100.14	9.1225	s
$4d^{10}12p$	$12p\,^2P^\circ_{\frac{1}{2}}$	*131261.1*	9.3650	
$4d^{10}12p$	$12p\,^2P^\circ_{1\frac{1}{2}}$	*131374.7*	9.3696	

TABLE II.—*Continued.*

1	2	3	4	5
$4d^95s5p$	$y\,^2F^\circ_{2\frac{1}{2}}$	*131718.2*		
$4d^{10}13s$	$13s\,^2S_{\frac{1}{2}}$	131899.07	9.9034	
$4d^{10}10f$	$10f\,^2F^\circ_{3\frac{1}{2}}$	*131953.72?*	9.9640	
$4d^{10}10f$	$10f\,^2F^\circ_{2\frac{1}{2}}$	*131961.87?*	9.9734	
$4d^{10}10g$	$10g\,^2G$	131978.47	9.9922	
$4d^{10}12d$	$12d\,^2D_{1\frac{1}{2}}$	132086.77	10.1175	
$4d^{10}12d$	$12d\,^2D_{2\frac{1}{2}}$	132092.18	10.1242	
$4d^{10}13p$	$13p\,^2P^\circ_{1\frac{1}{2}}$	*132280.4?*	10.3541	
$4d^{10}11f$	$11f\,^2F^\circ_{3\frac{1}{2}}$	*132707.91?*	10.9410	
$4d^{10}11f$	$11g\,^2G$	132742.26	10.9928	
$4d^{10}13d$	$13d\,^2D_{1\frac{1}{2}}$	132824.25	11.119	
$4d^{10}13d$	$13d\,^2D_{2\frac{1}{2}}$	132828.29	11.125	
$4d^{10}14d$	$14d\,^2D_{2\frac{1}{2}}$	133389.76	12.126	

Cd III $4d^{10}\ {}^1S_0$ 136.374.74

Column 1. Electron structure. 2. Level designation. 3. Level value. 4. Rydberg denominator. 5. Previous observer: s = von Salis, t = Takahashi.

term scale. That there is a change in the excitation at this point is shown by the intensities, but it is not as marked as it is in many other cases. The terms up to $10s^2S$ and $9d^2D$ which are just above the point, give strong series lines, the remainders of the series being composed of very weak lines. On the other hand, the $^2P^\circ$ and $^2F^\circ$ levels give strong lines much further along the series. This may indicate a difference in excitation of odd and even levels.

The details of the Cd II spectrum are given in the term table (Table II) and the line list (Table III). The numerical values of most of the terms are determined from wave-length measurements made with the large grating. Where they are given to two places of decimals they are suitable for predicting accurate wave numbers of far ultraviolet lines, as has been done in a number of other spectra. This means that most of the Cd II lines below 2000A can be calculated and used as standards of wave-length.

CADMIUM III

It was very surprising to us to find that the third spectrum of cadmium was excited in the hollow cathode with great intensity and even more completely than the second spectrum. The main features of the spectrum were given in 1928 by Gibbs and White and by McLennan, McLay, and Crawford[7] who found the terms belonging to the structures $4d^{10}$, $4d^95s$, $4d^95p$. From our line lists we soon discovered a number of new even terms above and combining with $4d^95p$. It became evident that we had found not only the series member $4d^96s\ {}^3D^1D$, but that the structure $4d^85s^2$ was also present. It is now, in fact, quite complete if we include the

[7] R. C. Gibbs and H. E. White, Phys. Rev. **31**, 776 (1928); J. C. McLennan, A. B. McLay, and M. F. Crawford, Trans. Roy. Soc. Can. **22**, 45 (1928).

TABLE III. List of identified lines of cadmium II.

1	2	3	4
8066.99	50	12392.80	$6s^3S_{1\frac{1}{2}}-6p^2P_{1\frac{1}{2}}^\circ$
7284.38	100	13724.24	$6d^2D_{2\frac{1}{2}}-6f^2F_{3\frac{1}{2}}^\circ$
7275.75	10	13740.50	$6d^2D_{2\frac{1}{2}}-6f^2F_{2\frac{1}{2}}^\circ$
7237.01	50	13814.06	$6d^2D_{1\frac{1}{2}}-6f^2F_{2\frac{1}{2}}^\circ$
6935.46	5 *h*	14414.68	$z^4F_{3\frac{1}{2}}^\circ-6g^2G$
6818.39	20	14662.72	$7p^2P_{1\frac{1}{2}}^\circ-10s^2S_{\frac{1}{2}}$
6759.189	100	14790.60	$6p^2P_{1\frac{1}{2}}^\circ-6d^2D_{1\frac{1}{2}}$
6725.780	500	14864.07	$6p^2P_{1\frac{1}{2}}^\circ-6d^2D_{2\frac{1}{2}}$
6684.161	10	14956.62	$z^4P_{\frac{1}{2}}^\circ-10s^2S_{\frac{1}{2}}$
6574.098	5	15207.02	$7p^2P_{1\frac{1}{2}}^\circ-9d^2D_{1\frac{1}{2}}$
6567.648	25	15221.94	$7p^2P_{1\frac{1}{2}}^\circ-9d^2D_{2\frac{1}{2}}$
6464.936	400	15463.79	$6p^2P_{\frac{1}{2}}^\circ-6d^2D_{1\frac{1}{2}}$
6449.456	15	15500.91	$z^4P_{\frac{1}{2}}^\circ-9d^2D_{1\frac{1}{2}}$
6359.982	500	15718.98	$4f^2F_{3\frac{1}{2}}^\circ-6g^2G$
6354.724	400	15731.98	$4f^2F_{3\frac{1}{2}}^\circ-6g^2G$
5880.220	50	17001.47	$6d^2D_{2\frac{1}{2}}-7f^2F_{3\frac{1}{2}}^\circ$
5868.502	10	17035.41	$6d^2D_{2\frac{1}{2}}-7f^2F_{2\frac{1}{2}}^\circ$
5843.305	40	17108.87	$6d^2D_{1\frac{1}{2}}-7f^2F_{2\frac{1}{2}}^\circ$
5802.801	10	17228.29	$7p^2P_{1\frac{1}{2}}^\circ-10d^2D_{2\frac{1}{2}}$
5708.77	10	17512.06	$z^4P_{\frac{1}{2}}^\circ-10d^2D_{1\frac{1}{2}}$
5568.36	2	17953.64	$7s^3S_{1\frac{1}{2}}-9p^2P_{1\frac{1}{2}}^\circ$
5471.25	2	18272.29	$z^4P_{1\frac{1}{2}}^\circ-9d^2D_{2\frac{1}{2}}$?
5449.41	1	18345.53	$7p^2P_{1\frac{1}{2}}^\circ-12s^2S_{\frac{1}{2}}$
5381.887	200	18575.69	$5d^2D_{2\frac{1}{2}}-4f^2F_{2\frac{1}{2}}^\circ$
5378.134	1000	18588.65	$5d^2D_{2\frac{1}{2}}-4f^2F_{3\frac{1}{2}}^\circ$
5337.484	1000	18730.22	$5d^2D_{1\frac{1}{2}}-4f^2F_{2\frac{1}{2}}^\circ$
5271.600	100	18964.31	$4f^2F_{3\frac{1}{2}}^\circ-7g^2G$
5268.007	100	18977.24	$4f^2F_{2\frac{1}{2}}^\circ-7g^2G$
5194.66	15	19245.19	$6d^2D_{1\frac{1}{2}}-8f^2F_{2\frac{1}{2}}^\circ$
5142.90	10	19438.88	$6d^2D_{2\frac{1}{2}}-8f^2F_{3\frac{1}{2}}^\circ$
5025.50	50	19892.99	$5d^2D_{2\frac{1}{2}}-z^4F_{3\frac{1}{2}}^\circ$
4918.85	20	20324.29	$5d^2D_{2\frac{1}{2}}-z^4F_{2\frac{1}{2}}^\circ$
4881.725	50	20478.86	$5d^2D_{1\frac{1}{2}}-z^4F_{2\frac{1}{2}}^\circ$
4834.64	10	20678.30	$6d^2D_{2\frac{1}{2}}-9f^2F_{3\frac{1}{2}}^\circ$
4828.52	5	20704.52	$6d^2D_{1\frac{1}{2}}-9f^2F_{2\frac{1}{2}}^\circ$
4744.693	30	21070.30	$4f^2F_{3\frac{1}{2}}^\circ-8g^2G$
4741.776	20	21083.27	$4f^2F_{2\frac{1}{2}}^\circ-8g^2G$
4605.71	2	21706.12	$6d^2D_{2\frac{1}{2}}-10f^2F_{3\frac{1}{2}}^\circ$
4588.45	1	21787.77	$6d^2D_{1\frac{1}{2}}-10f^2F_{2\frac{1}{2}}^\circ$
4535.16	3	22043.78	$7s^3S_{1\frac{1}{2}}-y^2P_{1\frac{1}{2}}^\circ$
4451.00	2	22460.58	$6d^2D_{2\frac{1}{2}}-11f^2F_{3\frac{1}{2}}^\circ$
4441.76	5	22507.30	$5d^2D_{1\frac{1}{2}}-z^4P_{\frac{1}{2}}^\circ$
4440.45	30	22513.94	$4f^2F_{3\frac{1}{2}}^\circ-9g^2G$
4437.91	20	22526.84	$4f^2F_{2\frac{1}{2}}^\circ-9g^2G$
4415.63	1000	22640.49	$5p^2P_{1\frac{1}{2}}^\circ-a^2D_{2\frac{1}{2}}$
4412.41	100	22657.01	$6p^2P_{1\frac{1}{2}}^\circ-8s^2S_{\frac{1}{2}}$
4384.52	5	22801.13	$5d^2D_{1\frac{1}{2}}-7p^2P_{1\frac{1}{2}}^\circ$
4362.68	1	22915.27	$7s^2S_{\frac{1}{2}}-11p^2P_{1\frac{1}{2}}^\circ$
4285.078	100	23330.26	$6p^2P_{\frac{1}{2}}^\circ-8s^2S_{\frac{1}{2}}$
4245.760	5 *h*	23546.30	$4f^2F_{3\frac{1}{2}}^\circ-10g^2G$
4243.428	5 *h*	23559.24	$4f^2F_{2\frac{1}{2}}^\circ-10g^2G$
4141.493	50	24139.10	$6p^2P_{1\frac{1}{2}}^\circ-7d^2D_{1\frac{1}{2}}$
4134.768	200	24178.36	$6p^2P_{1\frac{1}{2}}^\circ-7d^2D_{2\frac{1}{2}}$
4112.367	5 *h*	24310.06	$4f^2F_{2\frac{1}{2}}^\circ-11g^2G$
4110.169	5 *h*	24323.06	$4f^2F_{2\frac{1}{2}}^\circ-11g^2G$
4044.830	1	24715.96	$6p^2P_{1\frac{1}{2}}^\circ-11d^2D_{2\frac{1}{2}}$?
4029.124	100	24812.30	$6p^2P_{\frac{1}{2}}^\circ-7d^2D_{1\frac{1}{2}}$
4006.867	20	24950.13	$5d^2D_{2\frac{1}{2}}-z^4D_{3\frac{1}{2}}^\circ$?
3957.244	15	25262.99	$5d^2D_{1\frac{1}{2}}-z^4D_{2\frac{1}{2}}^\circ$
3826.710	15	26124.72	$a^2D_{2\frac{1}{2}}-6p^2P_{1\frac{1}{2}}^\circ$
3779.630	1	26450.13	$6s^2S_{\frac{1}{2}}-z^4P_{1\frac{1}{2}}^\circ$
3767.336	5	26536.45	$5d^2D_{1\frac{1}{2}}-z^4D_{1\frac{1}{2}}^\circ$
3666.756	20	27264.32	$5d^2D_{1\frac{1}{2}}-z^2F_{2\frac{1}{2}}^\circ$
3535.690	100	28274.97	$5p^2P_{1\frac{1}{2}}^\circ-a^2D_{1\frac{1}{2}}$
3524.112	100	28367.86	$6p^2P_{1\frac{1}{2}}^\circ-9s^2S_{\frac{1}{2}}$

TABLE III.—*Continued.*

1	2	3	4
3495.436	150	28600.58	$5d^2D_{2\frac{1}{2}}-5f^2F_{3\frac{1}{2}}^\circ$
3483.079	25	28701.95	$5d^2D_{2\frac{1}{2}}-5f^2F_{2\frac{1}{2}}^\circ$
3464.426	100	28856.57	$5d^2D_{1\frac{1}{2}}-5f^2F_{2\frac{1}{2}}^\circ$
3442.416	50	29041.07	$6p^2P_{\frac{1}{2}}^\circ-9s^2S_{\frac{1}{2}}$
3422.998	8	29205.81	$6s^3S_{\frac{1}{2}}-z^4P_{\frac{1}{2}}^\circ$
3422.228	5	29212.38	$5d^2D_{2\frac{1}{2}}-z^2P_{1\frac{1}{2}}^\circ$
3420.194	20	29229.76	$6p^2P_{1\frac{1}{2}}^\circ-8d^2D_{1\frac{1}{2}}$
3417.491	50	29252.87	$6p^2P_{1\frac{1}{2}}^\circ-8d^2D_{2\frac{1}{2}}$
3388.884	30	29499.80	$6s^2S_{\frac{1}{2}}-7p^2P_{1\frac{1}{2}}^\circ$
3385.486	50	29529.41	$5d^2D_{2\frac{1}{2}}-z^2F_{3\frac{1}{2}}^\circ$
3376.866	10	29604.78	$5d^2D_{1\frac{1}{2}}-z^2P_{\frac{1}{2}}^\circ$
3355.362	10	29794.51	$6s^2S_{\frac{1}{2}}-z^4F_{1\frac{1}{2}}^\circ$
3343.209	50	29902.81	$6p^2P_{\frac{1}{2}}^\circ-8d^2D_{1\frac{1}{2}}$
3283.565	3	30445.96	$5d^2D_{1\frac{1}{2}}-z^2D_{1\frac{1}{2}}^\circ$
3250.328	150	30757.28	$5p^2P_{\frac{1}{2}}^\circ-m^2D_{1\frac{1}{2}}$
3238.742	10	30867.31	$5d^2D_{2\frac{1}{2}}-8p^2P_{1\frac{1}{2}}^\circ$
3232.260	5	30929.21	$5d^2D_{1\frac{1}{2}}-8p^2P_{\frac{1}{2}}^\circ$
3222.614	2	31021.78	$5d^2D_{1\frac{1}{2}}-8p^2P_{1\frac{1}{2}}^\circ$
3194.42	1	31295.54	$a^2D_{1\frac{1}{2}}-z^4P_{2\frac{1}{2}}^\circ$
3179.961	10	31437.86	$5d^2D_{2\frac{1}{2}}-z^2D_{2\frac{1}{2}}^\circ$
3164.343	1	31593.02	$5d^2D_{1\frac{1}{2}}-z^2D_{2\frac{1}{2}}^\circ$
3146.790	50	31769.24	$6p^2P_{1\frac{1}{2}}^\circ-10s^2S_{\frac{1}{2}}$
3093.769	20	32313.69	$6p^2P_{1\frac{1}{2}}^\circ-9d^2D_{1\frac{1}{2}}$
3092.337	100	32328.65	$6p^2P_{1\frac{1}{2}}^\circ-9d^2D_{2\frac{1}{2}}$
3081.484	25	32442.51	$6p^2P_{\frac{1}{2}}^\circ-10s^2S_{\frac{1}{2}}$
3030.605	50	32987.14	$6p^2P_{\frac{1}{2}}^\circ-9d^2D_{1\frac{1}{2}}$
2943.831	15	33959.44	$6p^2P_{1\frac{1}{2}}^\circ-11s^2S_{\frac{1}{2}}$
2929.271	200	34128.23	$5d^2D_{2\frac{1}{2}}-6f^2F_{3\frac{1}{2}}^\circ$
2927.867	50	34144.59	$5d^2D_{2\frac{1}{2}}-6f^2F_{2\frac{1}{2}}^\circ$
2914.672	200	34299.16	$5d^2D_{1\frac{1}{2}}-6f^2F_{2\frac{1}{2}}^\circ$
2912.491	5	34324.84	$6p^2P_{1\frac{1}{2}}^\circ-10d^2D_{1\frac{1}{2}}$
2911.627	20	34335.03	$6p^2P_{1\frac{1}{2}}^\circ-10d^2D_{2\frac{1}{2}}$
2893.740	10	34547.25	$a^2D_{1\frac{1}{2}}-z^4P_{1\frac{1}{2}}^\circ$
2886.607	10	34632.62	$6p^2P_{\frac{1}{2}}^\circ-11s^2S_{\frac{1}{2}}$
2856.458	25	34998.13	$6p^2P_{\frac{1}{2}}^\circ-10d^2D_{1\frac{1}{2}}$ $6s^2S_{\frac{1}{2}}-z^4D_{\frac{1}{2}}^\circ$
2834.081	20	35274.45	$a^2D_{1\frac{1}{2}}-z^4F_{2\frac{1}{2}}^\circ$
2823.189	30	35410.53	$5d^2D_{2\frac{1}{2}}-9p^2P_{1\frac{1}{2}}^\circ$
2819.865	10	35452.28	$6p^2P_{1\frac{1}{2}}^\circ-12s^2S_{\frac{1}{2}}$
2813.389	5	35533.88	$5d^2D_{1\frac{1}{2}}-9p^2P_{\frac{1}{2}}^\circ$
2810.894	0	35565.42	$5d^2D_{1\frac{1}{2}}-9p^2P_{1\frac{1}{2}}^\circ$
2799.570	1	35709.27	$6p^2P_{1\frac{1}{2}}^\circ-11d^2D_{1\frac{1}{2}}$
2799.000	20	35716.54	$6p^2P_{1\frac{1}{2}}^\circ-11d^2D_{2\frac{1}{2}}$
2771.923	20	36065.41	$6s^2S_{\frac{1}{2}}-z^2P_{1\frac{1}{2}}^\circ$
2767.320	2	36125.40	$6p^2P_{\frac{1}{2}}^\circ-12s^2S_{\frac{1}{2}}$
2753.759	20	36303.29	$6s^2S_{\frac{1}{2}}-z^2P_{\frac{1}{2}}^\circ$
2748.549	1000	36372.10	$5p^2P_{1\frac{1}{2}}^\circ-6s^2S_{\frac{1}{2}}$
2723.363	20	36708.55	$6p^2P_{1\frac{1}{2}}^\circ-12d^2D_{2\frac{1}{2}}$
2707.003	50	36930.29	$a^2D_{2\frac{1}{2}}-z^4P_{2\frac{1}{2}}^\circ$
2691.386	1	37144.57	$6s^2S_{\frac{1}{2}}-z^2D_{1\frac{1}{2}}^\circ$
2688.195	0	37188.66	$6p^2P_{\frac{1}{2}}^\circ-13s^2S_{\frac{1}{2}}$
2679.968	10	37302.81	$a^2D_{1\frac{1}{2}}-z^4P_{\frac{1}{2}}^\circ$
2674.694	4	37376.36	$6p^2P_{\frac{1}{2}}^\circ-12d^2D_{1\frac{1}{2}}$
2672.624	50	37405.31	$5d^2D_{2\frac{1}{2}}-7f^2F_{3\frac{1}{2}}^\circ$
2670.202	15	37439.29	$5d^2D_{2\frac{1}{2}}-7f^2F_{2\frac{1}{2}}^\circ$
2669.815	1	37444.66	$6p^2P_{1\frac{1}{2}}^\circ-13d^2D_{2\frac{1}{2}}$
2668.195	25	37467.40	$a^2D_{1\frac{1}{2}}-7p^2P_{\frac{1}{2}}^\circ$
2659.226	40	37593.76	$5d^2D_{1\frac{1}{2}}-7f^2F_{2\frac{1}{2}}^\circ$
2659.021	10	37596.66	$a^2D_{1\frac{1}{2}}-7p^2P_{1\frac{1}{2}}^\circ$
2650.293	8	37720.46	$6s^2S_{\frac{1}{2}}-8p^2P_{1\frac{1}{2}}^\circ$
2638.326	3	37891.54	$a^2D_{1\frac{1}{2}}-z^4F_{1\frac{1}{2}}^\circ$
2630.371	0	38006.13	$6p^2P_{1\frac{1}{2}}^\circ-14d^2D_{2\frac{1}{2}}$
2622.937	0	38113.84	$6p^2P_{\frac{1}{2}}^\circ-13d^2D_{1\frac{1}{2}}$
2572.930	500	38854.57	$5p^2P_{\frac{1}{2}}^\circ-6s^2S_{\frac{1}{2}}$
2552.827	15	39160.52	$a^2D_{2\frac{1}{2}}-4f^2F_{3\frac{1}{2}}^\circ$

somewhat doubtful level a^1S_0. There also appears a large group of levels overlapping $4d^96s$. They are due to $4d^95d$. All of these groups of levels give lines of distinctive character. The lines $4d^95p-4d^85s^2$ are sharp and relatively weak in the spark whereas $4d^96s$ and $4d^95d$ give strong diffuse lines in the spark. It was on that basis that the one remaining line of correct character was picked as the combination $4d^95p\ {}^1P_1^\circ-4d^85s^2\ {}^1S_0$ to determine a^1S_0.

TABLE III.—*Continued.*

1	2	3	4
2551.976	50	39173.58	$a^2D_{2\frac{1}{2}}-4f^2F_{3\frac{1}{2}}^\circ$
2530.815	15	39501.09	$5d^2D_{2\frac{1}{2}}-y^2P_{1\frac{1}{2}}^\circ$
2526.047?	2	39572.95	$5d^2D_{2\frac{1}{2}}-8f^2F_{2\frac{1}{2}}^\circ$?
2520.954	0	39655.60	$5d^2D_{1\frac{1}{2}}-y^2P_{1\frac{1}{2}}^\circ$
2516.224	30	39730.13	$5d^2D_{1\frac{1}{2}}-8f^2F_{2\frac{1}{2}}^\circ$
2509.107	50	39842.82	$5d^2D_{2\frac{1}{2}}-8f^2F_{3\frac{1}{2}}^\circ$
2495.584	40	40058.70	$a^2D_{1\frac{1}{2}}-z^1D_{2\frac{1}{2}}^\circ$
2487.918	40	40181.96	$a^2D_{2\frac{1}{2}}-z^1P_{1\frac{1}{2}}^\circ$
2476.179	2	40372.60	$5d^2D_{2\frac{1}{2}}-11p^2P_{1\frac{1}{2}}^\circ$
2469.733	50	40477.97	$a^2D_{2\frac{1}{2}}-z^4F_{3\frac{1}{2}}^\circ$
2443.696	4	40909.22	$a^2D_{2\frac{1}{2}}-z^4F_{2\frac{1}{2}}^\circ$
2433.415	10	41082.04	$5d^2D_{2\frac{1}{2}}-9f^2F_{3\frac{1}{2}}^\circ$
2427.074	5	41189.36	$5d^2D_{1\frac{1}{2}}-9f^2F_{2\frac{1}{2}}^\circ$
2418.694	50	41332.06	$a^2D_{1\frac{1}{2}}-z^1D_{1\frac{1}{2}}^\circ$
2387.326	2	41875.09	$5d^2D_{2\frac{1}{2}}-y^2F_{2\frac{1}{2}}^\circ$
2376.820	40	42060.17	$a^2D_{1\frac{1}{2}}-z^2F_{2\frac{1}{2}}^\circ$
2367.130	5	42232.33	$6s^2S_{\frac{1}{2}}-9p^2P_{1\frac{1}{2}}^\circ$
2365.361	5	42263.91	$\begin{cases}6s^2S_{\frac{1}{2}}-9p^2P_{1\frac{1}{2}}^\circ\\6d^2D_{1\frac{1}{2}}-10f^2F_{2\frac{1}{2}}^\circ\end{cases}$
2321.074	200	43070.25	$5p^2P_{1\frac{1}{2}}^\circ-5d^2D_{1\frac{1}{2}}$
2312.766	1000	43224.95	$5p^2P_{1\frac{1}{2}}^\circ-5d^2D_{2\frac{1}{2}}$
2312.42?	2?	43231.05	$a^2D_{2\frac{1}{2}}-7p^2P_{1\frac{1}{2}}^\circ$
2296.76?	1	43525.97	$a^2D_{2\frac{1}{2}}-z^4F_{1\frac{1}{2}}^\circ$
2290.110	20	43652.54	$a^2D_{1\frac{1}{2}}-5f^2F_{2\frac{1}{2}}^\circ$
2265.018	1000	44136.08	$5s^2S_{\frac{1}{2}}-5p^2P_{\frac{1}{2}}^\circ$
2263.649	5	44162.77	$a^2D_{1\frac{1}{2}}-z^2P_{1\frac{1}{2}}^\circ$
2251.54	5	44400.25	$a^2D_{1\frac{1}{2}}-z^2P_{\frac{1}{2}}^\circ$
2221.229	10	45006.07	$6s^2S_{\frac{1}{2}}-10p^2P_{1\frac{1}{2}}^\circ$
2209.653	20	45241.82	$a^2D_{1\frac{1}{2}}-z^2D_{1\frac{1}{2}}^\circ$
2194.557	1000	45553.00	$5p^2P_{\frac{1}{2}}^\circ-5d^2D_{1\frac{1}{2}}$
2187.794	100	45693.80	$a^2D_{2\frac{1}{2}}-z^4D_{2\frac{1}{2}}^\circ$
2186.300	15	45725.03	$a^2D_{1\frac{1}{2}}-8p^2P_{\frac{1}{2}}^\circ$
2181.880	5	45817.64	$a^2D_{1\frac{1}{2}}-8p^2P_{1\frac{1}{2}}^\circ$
2156.636	5	46353.89	$6s^2S_{\frac{1}{2}}-y^2P_{1\frac{1}{2}}^\circ$
2155.062	50	46387.74	$a^2D_{1\frac{1}{2}}-z^2D_{2\frac{1}{2}}^\circ$
2144.408	1000 *R*	46618.18	$5s^2S_{\frac{1}{2}}-5p^2P_{1\frac{1}{2}}^\circ$
2128.488	20	46966.82	$\begin{cases}6s^2S_{\frac{1}{2}}-11p^2P_{\frac{1}{2}}^\circ\\a^2D_{2\frac{1}{2}}-z^4D_{1\frac{1}{2}}^\circ\end{cases}$
λ (Vac)			
2117.48	5	47226.0	$6s^2S_{\frac{1}{2}}-11p^2P_{1\frac{1}{2}}^\circ$
2096.67	150	47694.7	$a^2D_{2\frac{1}{2}}-z^2F_{2\frac{1}{2}}^\circ$
2036.88	75	49094.6	$a^2D_{1\frac{1}{2}}-6f^2F_{2\frac{1}{2}}^\circ$
2033.10	50	49186.0	$a^2D_{2\frac{1}{2}}-5f^2F_{3\frac{1}{2}}^\circ$
2028.92	20	49287.4	$a^2D_{2\frac{1}{2}}-5f^2F_{2\frac{1}{2}}^\circ$
2008.14	100	49797.3	$a^2D_{2\frac{1}{2}}-z^2P_{1\frac{1}{2}}^\circ$
1995.43	200	50114.5	$a^2D_{2\frac{1}{2}}-z^2F_{3\frac{1}{2}}^\circ$
1986.89	30	50329.9	$a^2D_{1\frac{1}{2}}-9p^2P_{\frac{1}{2}}^\circ$
1985.65	15	50361.4	$a^2D_{1\frac{1}{2}}-9p^2P_{1\frac{1}{2}}^\circ$
1965.54	40	50876.6	$a^2D_{2\frac{1}{2}}-z^2D_{1\frac{1}{2}}^\circ$
1943.54	100	51452.4	$a^2D_{2\frac{1}{2}}-8p^2P_{1\frac{1}{2}}^\circ$
1922.23	300	52023.0	$a^2D_{2\frac{1}{2}}-z^2D_{2\frac{1}{2}}^\circ$
1908.76	15	52389.9	$a^2D_{1\frac{1}{2}}-7f^2F_{2\frac{1}{2}}^\circ$
1883.09	10	53104.2	$a^2D_{1\frac{1}{2}}-10p^2P_{1\frac{1}{2}}^\circ$
1880.31	10	53182.7	$a^2D_{1\frac{1}{2}}-10p^2P_{\frac{1}{2}}^\circ$
1836.48	12	54451.9	$a^2D_{1\frac{1}{2}}-y^2P_{1\frac{1}{2}}^\circ$
1833.98	10	54526.2	$a^2D_{1\frac{1}{2}}-8f^2F_{2\frac{1}{2}}^\circ$
1827.70	100	54713.7	$a^2D_{2\frac{1}{2}}-6f^2F_{3\frac{1}{2}}^\circ$
1827.14	10	54730.5	$a^2D_{2\frac{1}{2}}-6f^2F_{2\frac{1}{2}}^\circ$
1816.11	5	55062.7	$a^2D_{1\frac{1}{2}}-11p^2P_{\frac{1}{2}}^\circ$
1807.57	2	55322.6	$a^2D_{1\frac{1}{2}}-11p^2P_{1\frac{1}{2}}^\circ$
1785.84	100	55995.9	$a^2D_{2\frac{1}{2}}-9p^2P_{1\frac{1}{2}}^\circ$
1774.07	1	56367.5	$a^2D_{1\frac{1}{2}}-12p^2P_{\frac{1}{2}}^\circ$
1759.78	20	56825.4	$a^2D_{1\frac{1}{2}}-y^2D_{2\frac{1}{2}}^\circ$
1724.41	50	57990.8	$a^2D_{2\frac{1}{2}}-7f^2F_{3\frac{1}{2}}^\circ$

TABLE III.—*Continued.*

1	2	3	4
1723.39	2	58025.3	$a^2D_{2\frac{1}{2}}-7f^2F_{2\frac{1}{2}}^\circ$
1702.47	50	58738.4	$a^2D_{2\frac{1}{2}}-10p^2P_{1\frac{1}{2}}^\circ$
1668.60	100	59930.6	$a^2D_{2\frac{1}{2}}-y^2F_{3\frac{1}{2}}^\circ$
1664.27	60	60086.3	$a^2D_{2\frac{1}{2}}-y^2P_{1\frac{1}{2}}^\circ$
1662.22	1	60160.4	$a^2D_{2\frac{1}{2}}-8f^2F_{2\frac{1}{2}}^\circ$
1654.86	75	60428.0	$a^2D_{2\frac{1}{2}}-8f^2F_{3\frac{1}{2}}^\circ$
1647.93	60	60682.3	$5p^2P_{1\frac{1}{2}}^\circ-7s^2S_{\frac{1}{2}}$
1640.50	40	60957.2	$a^2D_{2\frac{1}{2}}-11p^2P_{1\frac{1}{2}}^\circ$
1622.84	5	61620.4	$a^2D_{2\frac{1}{2}}-9f^2F_{2\frac{1}{2}}^\circ$
1621.61	15	61667.1	$a^2D_{2\frac{1}{2}}-9f^2F_{3\frac{1}{2}}^\circ$
1609.90	10	62115.8	$a^2D_{2\frac{1}{2}}-12p^2P_{1\frac{1}{2}}^\circ$
1601.04	50	62459.6	$a^2D_{2\frac{1}{2}}-y^2F_{2\frac{1}{2}}^\circ$
1594.81	30	62703.4	$a^2D_{2\frac{1}{2}}-10f^2F_{2\frac{1}{2}}^\circ$
1586.76	1	63021.5	$a^2D_{2\frac{1}{2}}-13p^2P_{1\frac{1}{2}}^\circ$
1583.17	50	63164.6	$5p^2P_{\frac{1}{2}}^\circ-7s^2S_{\frac{1}{2}}$
1573.42	50	63555.9	$5p^2P_{1\frac{1}{2}}^\circ-6d^2D_{1\frac{1}{2}}$
1571.58	200	63630.3	$5p^2P_{1\frac{1}{2}}^\circ-6d^2D_{2\frac{1}{2}}$
1514.26	200	66038.7	$5p^2P_{\frac{1}{2}}^\circ-6d^2D_{1\frac{1}{2}}$
1404.11	75	71422.9	$5p^2P_{1\frac{1}{2}}^\circ-8s^2S_{\frac{1}{2}}$
1371.65	50	72904.7	$5p^2P_{1\frac{1}{2}}^\circ-7d^2D_{1\frac{1}{2}}$
1370.91	150	72944.1	$5p^2P_{1\frac{1}{2}}^\circ-7d^2D_{2\frac{1}{2}}$
1353.08	40	73905.7	$5p^2P_{\frac{1}{2}}^\circ-8s^2S_{\frac{1}{2}}$
1335.28	2	74890.9	$5s^2S_{\frac{1}{2}}-a^2D_{1\frac{1}{2}}$?
1326.50	100	75386.3	$5p^2P_{\frac{1}{2}}^\circ-7d^2D_{1\frac{1}{2}}$
1296.43	150	77134.8	$5p^2P_{1\frac{1}{2}}^\circ-9s^2S_{\frac{1}{2}}$
1282.12	15	77996.1	$5p^2P_{1\frac{1}{2}}^\circ-8d^2D_{1\frac{1}{2}}$
1281.73	75	78019.3	$5p^2P_{1\frac{1}{2}}^\circ-8d^2D_{2\frac{1}{2}}$
1256.00	100	79617.6	$5p^2P_{\frac{1}{2}}^\circ-9s^2S_{\frac{1}{2}}$
1242.54	75	80480.1	$5p^2P_{\frac{1}{2}}^\circ-8d^2D_{1\frac{1}{2}}$
1241.67	15	80536.6	$5p^2P_{1\frac{1}{2}}^\circ-10s^2S_{\frac{1}{2}}$
1233.32	5	81081.7	$5p^2P_{1\frac{1}{2}}^\circ-9d^2D_{1\frac{1}{2}}$
1233.10	20	81096.4	$5p^2P_{1\frac{1}{2}}^\circ-9d^2D_{2\frac{1}{2}}$
1208.82	3	82725.0	$5p^2P_{1\frac{1}{2}}^\circ-11s^2S_{\frac{1}{2}}$
1204.53	15	83019.7	$5p^2P_{\frac{1}{2}}^\circ-10s^2S_{\frac{1}{2}}$
1203.34	5	83102.3	$5p^2P_{1\frac{1}{2}}^\circ-10d^2D_{2\frac{1}{2}}$
1196.69	15	83563.6	$5p^2P_{\frac{1}{2}}^\circ-9d^2D_{1\frac{1}{2}}$
1187.40	1	84217.1	$5p^2P_{1\frac{1}{2}}^\circ-12s^2S_{\frac{1}{2}}$
1183.67	2	84483.2	$5p^2P_{1\frac{1}{2}}^\circ-11d^2D_{2\frac{1}{2}}$
1173.59	1	85208.4	$5p^2P_{\frac{1}{2}}^\circ-11s^2S_{\frac{1}{2}}$
1172.33	0	85300.3	$5p^2P_{1\frac{1}{2}}^\circ-13s^2S_{\frac{1}{2}}$?
1169.90	1	85477.2	$5p^2P_{1\frac{1}{2}}^\circ-12d^2D_{2\frac{1}{2}}$
1168.58	5	85574.0	$5p^2P_{\frac{1}{2}}^\circ-10d^2D_{1\frac{1}{2}}$
1153.38	1	86702.0	$5p^2P_{\frac{1}{2}}^\circ-12s^2S_{\frac{1}{2}}$
1149.98	1	86957.9	$5p^2P_{\frac{1}{2}}^\circ-11d^2D_{1\frac{1}{2}}$
1055.83	25	94712.5	$5s^2S_{\frac{1}{2}}-6p^2P_{\frac{1}{2}}^\circ$
1048.38	40	95385.3	$5s^2S_{\frac{1}{2}}-6p^2P_{1\frac{1}{2}}^\circ$
913.72	20	109442.7	$5s^2S_{\frac{1}{2}}-z^1P_{1\frac{1}{2}}^\circ$
891.28	25	112197.8	$5s^2S_{\frac{1}{2}}-z^1P_{\frac{1}{2}}^\circ$
889.98	15	112362.2	$5s^2S_{\frac{1}{2}}-7p^2P_{\frac{1}{2}}^\circ$
888.95	25	112492.7	$5s^2S_{\frac{1}{2}}-7p^2P_{1\frac{1}{2}}^\circ$
886.62	0	112787.5	$5s^2S_{\frac{1}{2}}-z^4F_{1\frac{1}{2}}^\circ$
860.38	30	116228.1	$5s^2S_{\frac{1}{2}}-z^4D_{1\frac{1}{2}}^\circ$
847.52	25	117991.4	$5s^2S_{\frac{1}{2}}-z^4D_{\frac{1}{2}}^\circ$
839.94	30	119056.6	$5s^2S_{\frac{1}{2}}-z^2P_{1\frac{1}{2}}^\circ$
838.26	30	119295.0	$5s^2S_{\frac{1}{2}}-z^2P_{\frac{1}{2}}^\circ$
832.38	40	120137.0	$5s^2S_{\frac{1}{2}}-z^2D_{1\frac{1}{2}}^\circ$
829.05	15	120620.4	$5s^2S_{\frac{1}{2}}-8p^2P_{\frac{1}{2}}^\circ$
828.40	2	120714.9	$5s^2S_{\frac{1}{2}}-8p^2P_{1\frac{1}{2}}^\circ$
798.57	15	125224.3	$5s^2S_{\frac{1}{2}}-9p^2P_{\frac{1}{2}}^\circ$
798.37	50	125255.5	$5s^2S_{\frac{1}{2}}-9p^2P_{1\frac{1}{2}}^\circ$
781.26	25	127998.0	$5s^2S_{\frac{1}{2}}-10p^2P_{1\frac{1}{2}}^\circ$
780.80	5	128074.5	$5s^2S_{\frac{1}{2}}-10p^2P_{\frac{1}{2}}^\circ$
773.12	20	129346.8	$5s^2S_{\frac{1}{2}}-y^2P_{1\frac{1}{2}}^\circ$
769.45	1	129962.7	$5s^2S_{\frac{1}{2}}-11p^2P_{\frac{1}{2}}^\circ$
767.94	10	130219.3	$5s^2S_{\frac{1}{2}}-11p^2P_{1\frac{1}{2}}^\circ$
761.18	1	131375.6	$5s^2S_{\frac{1}{2}}-12p^2P_{1\frac{1}{2}}^\circ$

Column 1. Wave-length. First four λ by von Salis. Remainder λ by authors. 2. Intensity. 3. Wave number. 4. Identification.

The presence of two members of the series $4d^9ns$ allows the calculation of an ionization potential for Cd III. From the *s*-electron series in Cd I and Cd II it is possible to predict a reasonable value for the Ritz correction in the Cd III series and so obtain a limit which is better than that calculated from the two terms considered as a Rydberg series. The result is 304,690 cm^{-1} or 37.6 volts.

TABLE IV. Terms of cadmium III.

1	2	3	4
$4d^{10}$	$4d^1S_0$	00.0	*G*
$4d^9(^2D_{2\frac{1}{2}})5s$	$5s^3D_3$	80454.3	*G*
$4d^9(^2D_{2\frac{1}{2}})5s$	$5s^3D_2$	82354.6	*G*
$4d^9(^2D_{1\frac{1}{2}})5s$	$5s^3D_1$	86219.5	*G*
$4d^9(^2D_{1\frac{1}{2}})5s$	$5s^1D_2$	88871.8	*G*
$4d^85p$	$5p^3P_2^\circ$	*133813.3*	*G*
$4d^85p$	$5p^3F_3^\circ$	*136214.3*	*G*
$4d^85p$	$5p^3P_1^\circ$	*138754.4*	*G*
$4d^85p$	$5p^3F_4^\circ$	*139031.2*	*G*
$4d^85p$	$5p^3D_2^\circ$	*140429.4*	*G*
$4d^85p$	$5p^3P_0^\circ$	*142110.8*	*G*
$4d^85p$	$5p^3F_2^\circ$	*142754.8*	*G*
$4d^85p$	$5p^3D_3^\circ$	*142892.8*	*G*
$4d^85p$	$5p^1P_1^\circ$	*146075.1*	*G*
$4d^85p$	$5p^1F_3^\circ$	*146090.6*	*G*
$4d^85p$	$5p^3D_1^\circ$	*147624.8*	*G*
$4d^85p$	$5p^1D_2^\circ$	*148461.3*	*G*
$4d^85s^2$	a^3F_4	179022.4	
$4d^85s^2$	a^3F_3	184094.3	
$4d^85s^2$	a^3F_2	185798.5	
$4d^85s^2$	a^3P_2	191762.3	
$4d^85s^2$	a^3P_1	196123.7	
$4d^85s^2$	a^3P_0	196258.2	
$4d^85s^2$	a^1D_2	198431.0	
$4d^85s^2$	a^1G_4	200932.8	
$4d^9(^2D_{2\frac{1}{2}})5d$	$5d^3S_1$	202966.1	
$4d^9(^2D_{2\frac{1}{2}})6s$	$6s^3D_3$	204421.1	
$4d^9(^2D_{2\frac{1}{2}})6s$	$6s^3D_2$	204908.1	
$4d^9(^2D_{2\frac{1}{2}})5d$	$5d^3G_5$	205119.4?	
$4d^9(^2D_{2\frac{1}{2}})5d$	$5d^3G_4$	205296.7	
$4d^9(^2D_{2\frac{1}{2}})5d$	$5d^3P_2$	205407.0	
$4d^9(^2D_{2\frac{1}{2}})5d$	$5d^3D_3$	206087.7	
$4d^9(^2D_{2\frac{1}{2}})5d$	$5d^3F_3$	206610.3	
$4d^9(^2D_{2\frac{1}{2}})5d$	$5d^3D_2$	206804.3	
$4d^9(^2D_{2\frac{1}{2}})5d$	$5d^3F_4$	206967.8	
$4d^9(^2D_{1\frac{1}{2}})5d$	$5d^1P_1$	210147.1	
$4d^9(^2D_{1\frac{1}{2}})6s$	$6s^3D_1$	210220.5	
$4d^9(^2D_{1\frac{1}{2}})6s$	$6s^1D_2$	210588.0	
$4d^9(^2D_{1\frac{1}{2}})5d$	$5d^3G_3$	210837.5	
$4d^9(^2D_{1\frac{1}{2}})5d$	$5d^1G_4$	211518.5	
$4d^9(^2D_{1\frac{1}{2}})5d$	$5d^3F_2$	211924.8	
$4d^9(^2D_{1\frac{1}{2}})5d$	$5d^3D_1$	212049.8	
$4d^9(^2D_{1\frac{1}{2}})5d$	$5d^1D_2$	212625.0	
$4d^9(^2D_{1\frac{1}{2}})5d$	$5d^1F_3$	212708.2	
$4d^85s^2$	a^1S_0	213107.7	
$4d^9\ ^2D_{2\frac{1}{2}}$Cd IV 304,690			

Column 1. Electron structure. 2. Name of level. 3. Value of level. 4. Previous observers.

G {Gibbs and White, Phys. Rev. **31**, 776 (1928). McLennan, McLay, Crawford, Trans. Roy. Soc. Can. **22**, 45 (1928).

TABLE V. List of identified lines of cadmium III.

1	2	3	4
3035.719	10	32931.57	$5p^1F_3^\circ - a^3F_4$
2805.588	30	35632.68	$5p^1D_2^\circ - a^3F_3$
2766.992	20	36129.68	$5p^3D_3^\circ - a^3F_4$
2630.558	5	38003.43	$5p^1F_3^\circ - a^3F_3$
2618.807	15	38173.95	$5p^3D_1^\circ - a^3F_2$
2499.81	25	39991.0	$5p^3F_4^\circ - a^3F_4$
2426.355	10	41201.8	$5p^3D_3^\circ - a^3F_3$
2418.243	7	41339.7	$5p^3F_2^\circ - a^3F_3$
2322.50	1	43043.8	$5p^3F_2^\circ - a^3F_2$
λ (Vac)			
2225.12	5	44941.4	$5s^1D_2^\circ - 5p^3P_2$
2219.12	5	45062.9	$5p^3F_4^\circ - a^3F_3$
2188.81	5	45686.9	$5p^1P_1^\circ - a^3P_2$
2112.27	50	47342.4	$5s^1D_2 - 5p^3F_3^\circ$
2101.14	5	47593.3	$5s^3D_1 - 5p^3P_2^\circ$
2098.12	10	47661.7	$5p^1D_2^\circ - a^3P_1$
2088.58	75	47879.5	$5p^3F_2^\circ - a^3F_3$
2061.91	10	48498.8	$5p^3D_1^\circ - a^3P_1$
2056.22	3	48633.0	$5p^3D_1^\circ - a^3P_0$
2046.26	50	48869.5	$5p^3D_3^\circ - a^3P_2$
2040.49	40	49007.9	$5p^3F_2^\circ - a^3P_2$
2016.78	15	49584.1	$5p^3F_3^\circ - a^3F_2$
2004.72	15	49882.2	$5s^1D_2 - 5p^3P_1^\circ$
2001.24	20	49968.9	$5p^1D_2^\circ - a^1D_2$
1992.72	1	50182.7	$5p^1P_1^\circ - a^3P_0$
1988.81	15	50281.4	$5p^3P_2^\circ - a^3F_3$
1968.25	3	50806.4	$5p^3D_1^\circ - a^1D_2$
1939.59	10	51557.2	$5s^1D_2 - 5p^3D_2^\circ$
1910.57	15	52340.4	$5p^1F_3^\circ - a^1D_2$
1909.98	25	52356.5	$5p^1P_1^\circ - a^1D_2$
1903.48	15	52535.5	$5s^3D_1 - 5p^3P_1^\circ$
1886.49	15	53008.5	$5p^3P_1^\circ - a^3P_2$
1874.08	150	53359.7	$5s^3D_3 - 5p^3P_2^\circ$
1856.67	200	53859.9	$5s^3D_2 - 5p^3F_3^\circ$
1855.85	40	53883.6	$5s^1D_2 - 5p^3F_2^\circ$
1851.37	20	54014.1	$5p^3P_0^\circ - a^3P_1$
1851.13	40	54021.2	$5s^1D_2 - 5p^3D_3^\circ$
1844.66	50	54201.6	$5s^3D_1 - 5p^3D_2^\circ$
1823.41	40	54842.4	$5p^1F_3^\circ - a^1G_4$
1800.57	5	55538.1	$5p^3D_3^\circ - a^1D_2$
1796.10	15	55676.63	$5p^3F_2^\circ - a^1D_2$
1793.40	75	55760.0	$5s^3D_2 - 5p^3F_3^\circ$
1789.19	30	55891.3	$5s^3D_1 - 5p^3P_0^\circ$
1773.06	40	56399.6	$5s^3D_2 - 5p^3P_1^\circ$
1768.82	30	56535.0	$5s^3D_1 - 5p^3F_2^\circ$

The analysis of Ag II made by E. Rasmussen[8] was of the greatest assistance in identifying the levels of $4d^85s^2$. Figure 3 shows how very similar

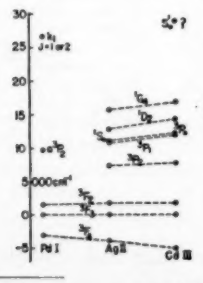

FIG. 3. $4d^85S^2$ structure in Pd I, Ag II, Cd III ($^3F_3 = 0$).

the structures are in the two spectra. If our identification of a^1S_0 is correct, the comparison should lead to a similar identification in Ag II. This would be of some importance, because such a level has never, we believe, been positively identified.

Although the levels of $4d^85s^2$ in Cd III and Ag II correspond in position and *J*-value, the choice of Russell-Saunders names for them depends somewhat on personal judgment because we are not really dealing with a case of *L-S* coupling. This accounts for the fact that the $^1S_0$ of Ag II corresponds to the $^3P_0$ of Cd III. The plot also shows the similar levels of Pd I.[9] The ^3F levels obviously correspond but the higher $J=2$ level is somewhat out of line. Pd I also has an unidentified level k_1, of $J=1$ or 2, which is shown on the diagram. It certainly does not appear to belong to the d^8s^2 group;

[8] E. Rasmussen, D. Kgl. Danske Vrdensk. Selskab, Math-Phys. Medd **XVIII**, 5 (1940).

[9] A. G. Shenstone, Phys. Rev. **36**, 669 (1930).

TABLE V.—*Continued.*

1	2	3	4
1748.15	15	57203.5	$5s^1D_2-5p^1P_1^\circ$
1747.67	40	57219.1	$5s^1D_2-5p^1F_3^\circ$
1745.69	5	57284.0	$5p^3D_1^\circ-6s^3D_2$
1743.10	2	57369.1	$5p^3P_1^\circ-a^3P_1$
1739.00	25	57504.4	$5p^3P_1^\circ-a^3P_0$
1725.66	30	57948.7	$5p^3P_2^\circ-a^3P_2$
1724.08	1	58001.9	$5p^3D_2^\circ-a^1D_2$
1722.95	40	58039.9	$5p^3D_3^\circ-a^1G_4$
1721.93	40	58074.5	$5s^3D_2-5p^3D_2^\circ$
1714.37	1	58330.6	$5p^1F_3^\circ-6s^3D_3$
1707.16	40	58576.9	$5s^3D_3-5p^3F_4^\circ$
1702.04	2	58753.0	$5s^1D_2-5p^3D_1^\circ$
1699.70	10	58833.8	$5p^1P_1^\circ-6s^3D_2$
1678.15	30	59589.5	$5s^1D_2-5p^1D_2^\circ$
1675.70	1	59676.5	$5p^3P_1^\circ-a^1D_2$
1670.69	1	59855.7	$5s^3D_1-5p^1P_1^\circ$
1666.77	0	59996.5	$5p^1F_3^\circ-5d^3D_3$
1655.63	25	60399.9	$5s^3D_2-5p^3F_2^\circ$
1651.87	20	60537.6	$5s^3D_2-5p^3D_3^\circ$
1628.54	25	61404.7	$5s^3D_1-5p^3D_1^\circ$
1625.27	15	61528.4	$5p^3D_3^\circ-6s^3D_3$
1621.13	3	61685.3	$5p^1D_2^\circ-5d^1P_1$
1619.21	3	61758.4	$5p^1D_2^\circ-6s^3D_1$
1615.47	3	61901.7	$5p^3F_4^\circ-a^1G_4$
1612.51	15	62015.0	$5p^3D_3^\circ-6s^3D_2$
1609.61	10	62126.7	$5p^1D_2^\circ-6s^1D_2$
1608.91	15	62154.0	$5p^3F_2^\circ-6s^3D_2$
1607.28	10	62216.8	$5p^3F_3^\circ-a^1D_2$
1606.64	20	62241.7	$5s^3D_1-5p^1D_2^\circ$
1604.87	20	62310.3	$5p^3P_2^\circ-a^3P_1$
1601.59	40	62438.0	$5s^3D_3-5p^3D_2^\circ$
1599.62	2	62514.8	$5p^3D_3^\circ-5d^3P_2$
1597.54	3	62596.4	$5p^3D_1^\circ-6s^3D_1$
1582.39	10	63195.5	$5p^3D_2^\circ-5d^3D_3$
1572.61	1	63588.5	$5p^1D_2^\circ-5d^3D_1$
1568.98	15	63735.8	$5s^3D_2-5p^1F_3^\circ$
1566.03	5	63855.8	$5p^3F_2^\circ-5d^3F_3$
1561.28	3	64050.0	$5p^3F_2^\circ-5d^3D_2$
1560.66	15	64075.3	$5p^3D_3^\circ-5d^3F_4$?
1557.35	3	64211.8	$5p^3P_1^\circ-5d^3S_1$
1556.48	5	64247.6	$5p^1D_2^\circ-5d^1F_3$
1555.21	3	64300.0	$5p^3D_2^\circ-5d^3F_2$?
1552.18	5	64425.5	$5p^3D_1^\circ-5d^3D_1$
1550.89	15	64478.9	$5p^3D_2^\circ-6s^3D_2$
1550.45	20	64497.2	$5p^1F_3^\circ-6s^1D_2$
1550.07	10	64513.1	$5p^1P_1^\circ-6s^1D_2$
1547.57	25	64617.5	$5p^3P_2^\circ-a^1D_2$

TABLE V.—*Continued.*

1	2	3	4
1545.17	50	64718.0	$5p^3F_3^\circ-a^1G_4$
1538.46	1	65000.2	$5p^3D_1^\circ-5d^1D_2$
1532.10	5	65269.9	$5s^3D_3-5p^3D_1^\circ$
1529.30	30	65389.6	$5p^3F_4^\circ-6s^3D_3$
1528.40	15	65427.9	$5p^1F_3^\circ-5d^1G_4$?
1523.55	10	65636.3	$5s^3D_3-5p^1F_3^\circ$
1513.13	5	66088.2	$5p^3F_4^\circ-5d^3G_5$?
1512.59	1	66111.8	$5s^3D_2-5p^1D_2^\circ$
1511.65	10	66153.1	$5p^3P_1^\circ-6s^3D_2$
1511.01	5	66181.1	$5p^3D_2^\circ-5d^3F_3$
1509.08	2	66265.7	$5p^3F_4^\circ-5d^3G_4$
1506.60	2	66374.5	$5p^3D_2^\circ-5d^3D_2$
1502.65	2	66549.1	$5p^1P_1^\circ-5d^1D_2$
1501.12	1	66616.8	$5p^1F_3^\circ-5d^1F_3$
1491.81	15	67032.6	$5p^1P_1^\circ-a^1S_0$?
1491.28	1	67056.4	$5p^3F_4^\circ-5d^3D_3$
1482.24	3	67465.4	$5p^3F_2^\circ-6s^3D_1$
1477.22	2	67694.8	$5p^3D_3^\circ-6s^1D_2$
1474.20	2	67833.3	$5p^3F_2^\circ-6s^1D_2$
1471.97	5	67936.2	$5p^3F_4^\circ-5d^3F_4$
1470.43	2	68007.2	$5s^3D_3-5p^1D_2^\circ$
1469.79	3	68036.7	$5p^3P_0^\circ-5d^1P_1$?
1469.51	1	68050.0	$5p^3P_1^\circ-5d^3D_2$
1468.80	2	68082.6	$5p^3F_2^\circ-5d^3G_3$
1468.22	2	68109.7	$5p^3P_0^\circ-6s^3D_1$
1466.14	8	68206.2	$5p^3F_3^\circ-6s^3D_3$
1455.74	30	68693.6	$5p^3F_3^\circ-6s^3D_2$
1447.55	25	69082.1	$5p^3F_3^\circ-5d^3G_4$
1446.08	20	69152.7	$5p^3P_2^\circ-5d^3S_1$
1445.23	1	69193.3	$5p^3F_2^\circ-5d^3P_2$
1432.86	5	69790.7	$5p^3D_2^\circ-6s^3D_1$
1431.17	3	69873.0	$5p^3F_3-5d^3D_3$
1429.82	2	69938.7	$5p^3P_0^\circ-5d^3D_1$
1425.35	1	70158.3	$5p^3D_2^\circ-6s^1D_2$
1420.54	8	70395.7	$5p^3F_3^\circ-5d^3F_3$
1420.29	5	70408.3	$5p^3D_2^\circ-5d^3G_3$
1416.28	25	70607.7	$5p^3P_2^\circ-6s^3D_3$
1406.59	2	71094.0	$5p^3P_2^\circ-6s^3D_2$
1400.70	3	71392.7	$5p^3P_1^\circ-5d^1P_1$
1399.27	1	71465.8	$5p^3P_1^\circ-6s^3D_1$
1396.78	10	71593.1	$5p^3P_2^\circ-5d^3P_2$
1392.10	15	71834.2	$5p^3P_1^\circ-6s^1D_2$
1385.11	1	72196.6	$5p^3D_2^\circ-5d^1D_2$
1383.60	5	72275.0	$5p^3P_2^\circ-5d^3D_3$
720.70	10	138754.5	$4d^1S_0-5p^3P_1^\circ$
684.58	15	146074.9	$4d^1S_0-5p^1P_1^\circ$
677.39	8	147625.0	$4d^1S_0-5p^3D_1^\circ$

Column 1. Wave-length. 2. Intensity. 3. Wave number. 4. Identification.

but, if it does not, then it is the only level in any spectrum which has no logical structural origin.

The levels of $4d^95d$ were mainly found from a comparison with the equivalent levels in Ag II.[10] The parallelism is not as striking as it is in the d^8s^2 structure, but there can be little doubt about the majority of the levels. The few which remain undiscovered have resisted a very careful search amongst the remaining lines of the correct character. The identified levels of Cd III are given in Table IV and the lines in Table V.

UNIDENTIFIED LINES

In the whole region observed with the hollow cathode discharge, there remain unidentified about 175 lines, most of which are of low intensity, a few only being of strength as high as 20 on our very open scale. The strongest lines include a number which are diffuse, indicating auto-ionization and a probable origin in Cd I, but no logical system of levels has been found to include them. Most of the remaining lines, especially two groups around λ3100 and λ1350 can be attributed to Cd III. Many of the lines have the characteristic of being strong in the spark but much less diffuse than the lines of $4d^95p-4d^95d$. They are probably a few of the very many lines which should appear from the combinations of $4d^95p$ and $4d^85s5p$ with lower terms. These two structures should exist at about the same level, but no convincing identifications of their combinations have been possible from the existing observations.

[10] A. G. Shenstone, Phys. Rev. **31**, 317 (1928).

Out of Focus Diffraction Patterns*

L. IVAN EPSTEIN
Scientific Bureau, Bausch & Lomb Optical Company, Rochester, New York
(Received November 5, 1948)

The diffraction pattern which represents the image of a luminous point produced by an instrument free from aberrations and having a circular aperture, when observed out of focus, has been calculated by A. E. Conrady, A. Buxton, and G. Lansraux. Conrady and Buxton used numerical integration. Such a process yields a sequence of values of the complex amplitude at a particular point in the diffraction pattern for a particular lack of focus, and for any given integration, the distance W of the point under consideration from the center of the pattern, and the lack of focus ψ, are functions of the upper limit of integration. If ψ is plotted against W, the path obtained is a parabola, $\psi = aW^2$.

The author has found a method whereby any desired path can be followed instead of a parabola. In particular, if the point of observation is near the edge of the purely geometrical diffusion disk and remains at a constant distance from it as ψ changes, then, for large values of ψ, the form of the diffraction fringes is approximately independent of ψ and their intensity is inversely proportional to ψ^2.

CONSIDER the image of a luminous point formed by an instrument free from aberrations and having a circular aperture. When the image is thrown out of focus, it broadens out into a disk which shows circular fringes in its interior. As the lack of focus increases, new fringes spring into view at the center of the field and expand outward. The outer fringes expand without showing any marked changes in appearance, i.e., the distance of a given fringe from the edge of the purely geometrical diffusion disk is approximately a constant independent of the lack of focus. It has been my aim to derive a mathematical formula for the distribution of light in the diffraction pattern from which the observed features would be readily apparent.

Let P (Fig. 1) be the focus towards which the spherical wave front A emerging from the instrument converges. Let B be another sphere, tangent to A at the apex, and having its center at Q, the center of the diffraction pattern in the plane of observation. For light of a given wave-length λ, the phase lag ψ between B and A at the edge of the exit pupil is proportional to the lack of focus. Let r be the angle between the direction of the diffracted light and the direct light through the center of the exit pupil, and let H be the radius of the exit pupil. Let $W = 2\pi Hr/\lambda$. The form of the diffraction pattern is then most easily described by the complex amplitude G, which is defined by the fact that its modulus gives the actual amplitude of the light and its argument gives the phase. Let $\Gamma = \lambda^2 G/\pi H^2 = \Gamma_1 + i\Gamma_2$. We can now describe the diffraction pattern in terms of the dimensionless quantities ψ, W, and Γ. We have

$$\Gamma = (2/W^2)\int_0^W \exp(-iv^2\psi/W^2) J_0(v) v dv = , \quad (1a)$$

$$= 2\int_0^1 \exp(-i\psi u^2) J_0(Wu) u du, \quad (1b)$$

where J is a Bessel function. This formula has been derived by a number of investigators. Conrady[1] and Buxton[2] evaluated Γ by numerical integration. Lansraux[3] gives the series expansion

$$\Gamma = \exp(-i\psi) \sum_{p=1}^{\infty} 2^p (i\psi)^{p-1} J_p(W)/W^p. \quad (2)$$

When ψ is large, the series (2) converges too slowly to be useful for numerical calculations, and if it is desired to evaluate Γ by numerical integration, this must be done in a large number of steps. Suppose we use formula (1a). During the numerical integration, the quantity ψ/W^2 in the argument of the exponential function is treated as a constant and will be denoted by a. Now, numerical integration is a step by step process, and at the end of each step, a particular value of the upper limit of integration W has been reached. To this value there corresponds a value of ψ which is given by the formula

$$\psi = aW^2. \quad (3)$$

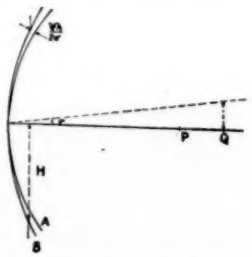

FIG. 1. Wave front emerging from the exit pupil of an instrument, and reference sphere.

* Presented at the meeting of the Optical Society of America in Detroit, Michigan, 1948.

[1] A. E. Conrady, M. N. R. A. S. **79**, 575 (1919)
[2] A. Buxton, M. N. R. A. S. **81**, 547 (1921).
[3] G. Lansraux, Revue d'Optique **26**, 24 (1947).

The integral in Eq. (1a) gives $\frac{1}{2}W^2\Gamma$. If this integral is evaluated by numerical integration, we will obtain $\frac{1}{2}W^2\Gamma$ for a sequence of values of ψ and W such that if ψ is plotted against W, the resulting graph is a parabola (Fig. 2). If it is desired to determine Γ for a given value of ψ and a variety of values of W, we must integrate along as many parabolic paths as values of W are desired. It would be much more convenient if we had a process of numerical integration by which we could follow the path $\psi=$const., since then a single numerical integration will give an entire diffraction pattern. Such a process may be obtained as follows:

Differentiating the series in (2) term by term and comparing the resulting series to the original one, we have

$$\partial\Gamma/\partial W=iW\Gamma/2\psi-i\exp(-i\psi)J_1(W)/\psi. \quad (4)$$

ψ is held constant in the differentiation. Therefore, the linear differential equation (4) may be integrated by elementary methods. For the case $W=0$, the integral (1b) can be expressed in terms of elementary functions[3]

$$\Gamma=(i/\psi)[\exp(-i\psi)-1], \quad (5)$$

and from this we get the constant of integration in the solution of Eq. (4). This solution is

$$\Gamma=\exp(iW^2/4\psi)\left\{(i/\psi)[\exp(-i\psi)-1]-\int_0^W(i/\psi)\exp(-i\psi-iW^2/4\psi)J_1(W)dW\right\}. \quad (6)$$

This is the desired formula, since ψ is to be held constant during the integration.

To explain the very simple changes in the diffraction pattern as ψ increases, we proceed as follows: Differentiating (1b) under the integral sign, we have

$$\partial\Gamma/\partial W=-2\int_0^1\exp(-i\psi u^2)J_1(Wu)u^2du, \quad (7)$$

$$\partial\Gamma/\partial\psi=-2i\int_0^1\exp(-i\psi u^2)J_0(Wu)u^3du. \quad (8)$$

The integral in (8) may be expressed differently by integration by parts. Comparing the resulting integral to (7) and (1b) we have

$$\partial\Gamma/\partial\psi=(1/\psi)\exp(-i\psi u^2)J_0(Wu)u^2\Big|_0^1-(1/\psi)\int_0^1\exp(-i\psi u^2)[2uJ_0(Wu)-Wu^2J_1(Wu)]du=$$
$$=(1/\psi)\exp(-i\psi)J_0(W)-\Gamma/\psi-(W/2\psi)(\partial\Gamma/\partial W). \quad (9a)$$

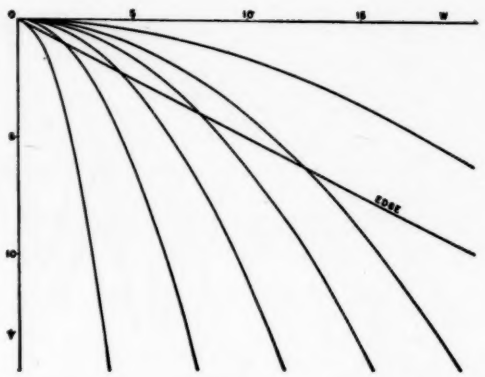

FIG. 2. Parabolic paths of integration and path followed by the edge of the purely geometrical diffusion disk.

Substituting from (4), we have

$$\partial\Gamma/\partial\psi=[\exp(-i\psi)/\psi][J_0(W)+(iW/2\psi)J_1(W)]-[1/\psi+iW^2/4\psi^2]\Gamma. \quad (9b)$$

Now, if W is any arbitrarily prescribed function of ψ, we have

$$d\Gamma/d\psi=\partial\Gamma/\partial\psi+(\partial\Gamma/\partial W)(dW/d\psi), \quad (10)$$

where $\partial\Gamma/\partial\psi$ and $\partial\Gamma/\partial W$ are known from (9b) and (4), respectively. Equation (10) is always a linear differential equation, which can be integrated by elementary methods. In particular, the edge of the purely geometrical diffusion disk is given by $W=2\psi$, and a point whose distance from the edge of the purely geometrical diffusion disk is independent of ψ is given by

$$W=2\psi+b. \quad (11)$$

From Eqs. (11), (10), (9b), and (4) we have

$$d\Gamma/d\psi=(i/2\psi^2)(-2\psi+b)\exp(-i\psi)J_1(W)+(1/\psi)\exp(-i\psi)J_0(W)+[i(4\psi^2-b^2)/4\psi^2-1/\psi]\Gamma. \quad (12a)$$

The integral of this equation is

$$\Gamma=(1/\psi)\exp(i\psi+ib^2/4\psi)\left\{C+(i/2-ib/4\psi)\times J_0(W)\exp(-2i\psi-ib^2/4\psi)+\int[b/2\psi-ib/4\psi^2+b^2/8\psi^2-b^3/16\psi^3]J_0(W)\times\exp(-2i\psi-ib^2/4\psi)d\psi\right\}, \quad (12b)$$

where W is given by formula (11). The constant of integration C is a function of the parameter b. If the lack of focus ψ is large and the distance $|b|$ of the point under consideration from the edge of

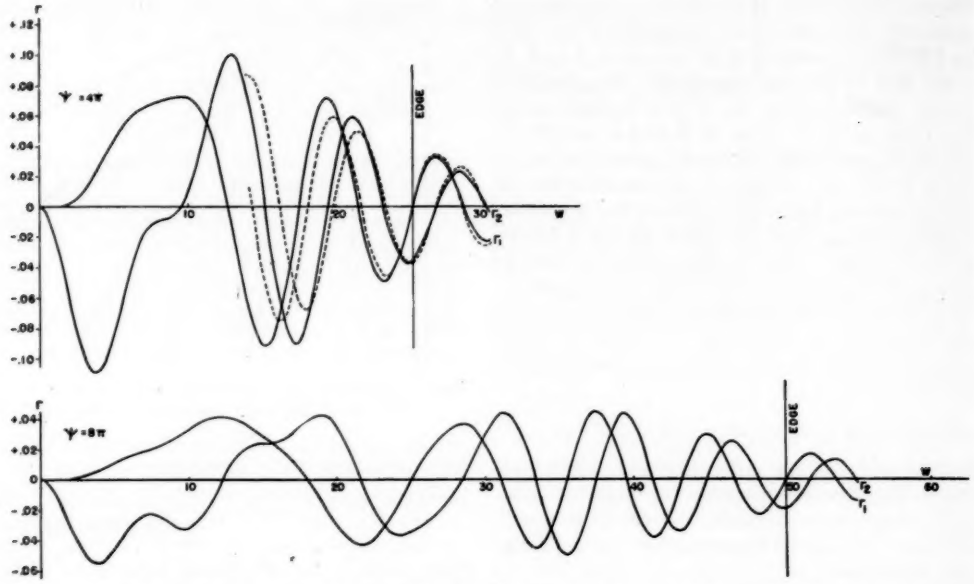

FIG. 3. Plot of Γ_1 and Γ_2 against W for $\psi=4\pi$ (above) and $\psi=8\pi$ (below). The dotted curves above give an approximation to the curves for $\psi=4\pi$ derived from those for $\psi=8\pi$ with the aid of formula (12d).

the purely geometrical diffusion disk is small, we may neglect all terms in b/ψ, and we have

$$\Gamma \doteq (1/\psi)\exp(i\psi)[C+\tfrac{1}{2}i\exp(-2i\psi)J_0(W)]. \quad (12c)$$

To obtain a general idea of the order of magnitude of the constant C, we may note that near $b=-9.818$, $|C|$ has a minimum of 2.313. Hence, for $W>10$, $|\frac{1}{2}J_0(W)|$ is at most about 5 percent of $|C|$. Therefore, when ψ is large and $|b|$ is small, the formula

$$\Gamma \doteq (1/\psi)\exp(i\psi)C \quad (12d)$$

gives a fair approximation.

Since the factor $\exp(i\psi)$ affects only the phase and not the intensity distribution in the diffraction pattern, it follows that as the image is thrown more and more out of focus, the outer diffraction fringes move outward at the same rate as the edge of the purely geometrical diffusion disk, and their relative intensity remains unchanged. At the same time, the intensity (which is proportional to the square of the amplitude) of the pattern as a whole is proportional to $1/\psi^2$, which was to be expected from geometrical optics.

Figure 3 shows the graphs of the real and imaginary component of Γ for $\psi=4\pi$ and $\psi=8\pi$, as well as (the dotted curves) an approximation to the former curves derived from the latter by formula (12d). The solid curves were obtained by numerical integration from formula (6).

The author wishes to thank Miss Ethel M. Thiel for help with the computations.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 39, NUMBER 3 MARCH, 1949

The Birefringent Filter

JOHN W. EVANS
Harvard Observatory, Harvard University, Cambridge, Massachusetts
(Received November 19, 1948)

Optical birefringent filters, which depend for their action on the interference of polarized light, can be designed to transmit very sharp bands (down to a fraction of an angstrom in width). The elementary theory necessary for their design is given.

Three forms of wide field filters designed by Lyot are described in detail. A more recently developed split element, wide field filter requires only half as many polarizers as the earlier types, which may be an advantage for some applications.

Various methods of adjusting the transmission bands of a birefringent filter, including the use of elements of variable thickness, and phase shifters are discussed. For most purposes the electro-optical phase shifters are probably the most promising. For special purposes, such as spectrophotometry, phase shifters composed of rotating fractional wave plates may be more advantageous. Two such phase shifters and their application in simple and split element filters are described.

A few crystalline materials which have been used or might be used to advantage in birefringent filters are mentioned.

Finally, the possibility of using polarizing interferometers in combination with birefringent elements for filters with extremely sharp transmission bands (in the range of hundredths or thousandths of an angstrom) is very briefly discussed.

I. INTRODUCTION

DURING the past few years the birefringent filter has proved an effective tool in astronomical research. Its utility, however, is not confined to astronomy and the purpose of the present paper is partly to bring it to the attention of investigators in other fields.

Briefly, the birefringent filter serves the purpose of a monochromator over an extended field. It can be designed to transmit a wave-length band of any desired width (down to a fraction of an angstrom) centered at any selected wave-length. It is used very much like an ordinary glass or gelatin filter in either a collimated or a converging beam of light, but with some limitation in field size or focal ratio, according to the type of construction, material, and band width.

The invention of the birefringent filter is one of the many important contributions of the French astronomer, Bernard Lyot,[1] to instrumental astronomy. He first published the basic principles of its operation in 1933. Öhman[2] independently invented the filter and in 1938 constructed the first one to be used for solar observations, with a transmission band about 40-angstroms wide centered on the $H\alpha$ line. With it he succeeded in seeing and photographing the brighter prominences, although it was evident that a much sharper band would be necessary for the best results.

In a later paper Lyot[3] has given a very complete discussion of the history, theory, and construction of birefringent filters. For the benefit of readers to whom his papers are not readily available, the present paper reviews enough of the elementary theory to suffice for the design of filters of any feasible characteristics. The remainder of the paper is a discussion of newer developments which serve to simplify the construction of the filters and extend their field of usefulness.

[1] B. Lyot, Comptes Rendus **197**, 1593 (1933).
[2] Y. Öhman, Nature **141**, No. 3560, 157 (1938); Nature **141**, No. 3563, 291 (1938); Populär Astronomisk Tidskrift, No. 1–2, pp. 11 and 27 (1938).
[3] B. Lyot, Ann. Astrophys. **7**, No. 1–2, p. 31 (1944).

II. THE SIMPLE BIREFRINGENT FILTER

Several forms of the birefringent filter are possible, differing in width of field and complexity of construction. They all depend, however, on the interference of polarized light transmitted through layers of birefringent crystal in the direction perpendicular to the plane of the optic axes if the crystal is biaxial, or any direction perpendicular to the optic axis if the crystal is uniaxial.

Since we can regard the uniaxial crystal as a degenerate biaxial crystal, most of the following discussions will consider only the biaxial case. Let ϵ and ω be the extraordinary and ordinary indices of refraction of any uniaxial crystal, and α, β, γ be the smallest, intermediate, and greatest principal indices of refraction of a biaxial crystal, respectively. Any expression for a biaxial crystal is valid for a uniaxial crystal if one of the following substitutions is made:

$$\alpha=\omega, \quad \beta=\omega, \quad \text{and} \quad \gamma=\epsilon \quad \text{if} \quad \epsilon-\omega>0,$$

or

$$\alpha=\epsilon, \quad \beta=\omega, \quad \text{and} \quad \gamma=\omega \quad \text{if} \quad \epsilon-\omega<0.$$

Unless otherwise specified, the mutually perpendicular directions of vibration of light for which the refractive indices are α, β, and γ will be referred to as the α-axis, β-axis, and γ-axis. These are, of course, the principal axes of the index ellipsoid.

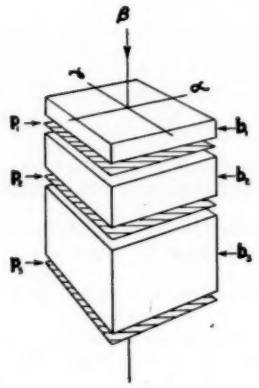

FIG. 1. Birefringent filter of three elements.

The quantity μ is defined by

$$\mu = \gamma - \alpha.$$

The term "retardation" will be used to indicate a path difference in terms of wave-lengths.

For brevity, the direction of vibration of light transmitted by a polarizer (prism or film) will be referred to as the axis of the polarizer.

Consider a block of some birefringent crystal, b_1 in Fig. 1, cut with its surfaces normal to its β-axis. Let light plane polarized at an angle of 45° to the α-axis enter the crystal along the β-axis. In the crystal the light divides into two components polarized with vibrations parallel to the α- and γ-axes, traveling with different velocities, c/α and c/γ. On emerging from the crystal, the two components have therefore a relative retardation of n_1, given by:

$$n_1 = (d_1/\lambda)\mu, \tag{II.1}$$

where d is the thickness of the crystal in the β-direction, and λ is the wave-length of the light.

If now the light traverses a polarizer, p_1 (which may be either a Nicol or similar prism, or a film polarizer), with its axis parallel to the vibration plane of the entering light, the two components interfere. The transmission, τ_1, of the b_1, p_1 combination is:

$$\tau_1 = \cos^2 \pi n_1. \tag{II.2}$$

If white light traverses the combination, the spectrum of the emergent light consists of regularly spaced alternate bright and dark bands at wave-lengths where n_1 is alternately integral and half-integral. The transmission as a function of wave-length is represented by curve a, Fig. 2.

The wave-length interval between successive bright bands is inversely proportional to the thickness of the crystal. To obtain an approximation of the interval, set $\Delta n = 1$ in the equation:

$$\frac{\Delta\lambda}{\lambda} = \frac{\Delta n}{n} \frac{1}{(\lambda/\mu)(\partial\mu/\partial\lambda) - 1}. \tag{II.3}$$

We now add a second crystal, b_2, and a polarizer p_2 oriented parallel to b_1 and p_1. If $d_2 = 2d_1$, the transmission of the b_2, p_2 combination, represented by curve b, Fig. 2, is:

$$\tau_2 = \cos^2 \pi n_2 = \cos^2 \pi 2n_1. \tag{II.4}$$

The transmission of the whole assembly, b_1, p_1, b_2, p_2, shown in curve c, Fig. 2, is therefore:

$$\tau_{12} = \cos^2 \pi n_1 \cos^2 \pi 2n_1. \tag{II.5}$$

A third crystal element, b_3, with $d_3 = 2d_2$, followed by the polarizer, p_3, has individual transmission shown in curve d. The transmission of the assembly, b_1 to p_3, is then represented by curve e, Fig. 2.

It is evident that further crystal elements and polarizers can be added. The result is the basic type of birefringent filter, which will be termed the simple filter. It is comprised of a series of units, each consisting of a plane-parallel birefringent element (b-element) followed by a polarizer. All b-elements have surfaces normal to their β-axes and are mounted with their α-axes parallel. All polarizers have their axes parallel to the vibration plane of the entering polarized light at 45° to the α-axes. The thickness of the rth b-element is such that

$$n_r = 2^{r-1} n_1. \tag{II.6}$$

The spectrum of light transmitted by the filter consists of a series of widely spaced narrow bands. Their separation is equal to the separation of the transmission maxima of the thinnest element alone,

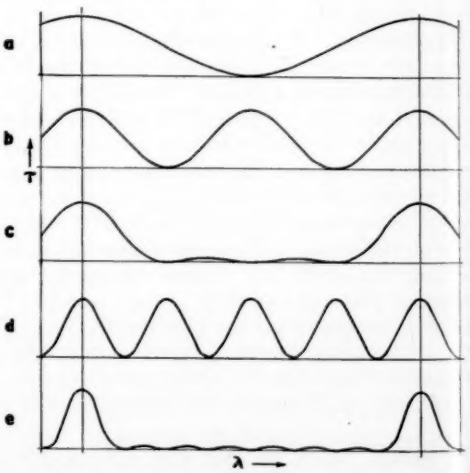

FIG. 2. Transmission curves for elements of Fig. 1. (a) b_1p_1; (b) b_2p_2; (c) $b_1p_1b_2p_2$; (d) b_3p_3; (e) $b_1p_1b_2p_2b_3p_3$.

while their effective width is the half-width of the maxima of the thickest element alone. For polarized entering light, the transmission of a filter of l b-elements (neglecting absorption in the material of the filter) is:

$$\tau = \cos^2\pi n_1 \cos^2\pi 2n_1 \cdots \cos^2\pi 2^{l-1} n_1. \quad \text{(II.7)}$$

The quantity n_1 must, of course, be an integer at the wave-length of the desired transmission band. Its magnitude should be small enough to separate the adjacent bands sufficiently to permit the isolation of the selected band by means of ordinary filters.

It can readily be shown that the total transmission of flux in an equal energy spectrum is 2^{-l}. Regardless of the width and separation of the bands, the total residual flux transmitted between successive principal maxima in a filter with $l>3$ is a substantially constant fraction (about 0.11) of the flux transmitted in a single band.

The filter at the Climax, Colorado station of the High Altitude Observatory of Harvard University and the University of Colorado has been in satisfactory operation in the observation of solar prominences since early 1943. It is a simple filter of six quartz elements with $n_1=23$, $n_6=736$, $d_1=1.677$ mm, and $d_6=53.658$ mm and has a transmission band of effective width 4.1 angstroms centered on the $H\alpha$ line of hydrogen (λ6563) at an operating temperature of 35.5°C. Its purpose is to eliminate the overpowering scattered light (continuous spectrum) near the limb of the sun while still transmitting the $H\alpha$-emission from the prominences, which are otherwise completely invisible. The success of the filter can be judged from the photographs in Plate I.

In practice, a filter should either be cemented or immersed in oil to avoid multiple reflections. Initial polarization is usually obtained by a polarizer mounted in front of the first b-element with its axis parallel to the axes of the other polarizers.

In any birefringent crystal, both the geometrical dimensions and μ are functions of temperature. The result is a small shift in the wave-lengths of the transmission maxima when the temperature changes. In quartz, for instance, $\Delta\lambda/\Delta T=-0.66$ angstrom per degree centigrade in the red. Hence the temperature of the filter must be controlled with sufficient accuracy to keep the maximum excursions of wave-length within tolerable limits. A total range of two-tenths of the effective band width is small enough for most purposes.

III. OFF-AXIS EFFECTS IN SIMPLE FILTERS

It is evident that when light traverses a simple filter at an angle to the instrumental axis, the light path through the birefringent material and the velocity difference of the fast and slow waves are altered. The effect is simply to alter the value of n_1 in Eq. (II.7).

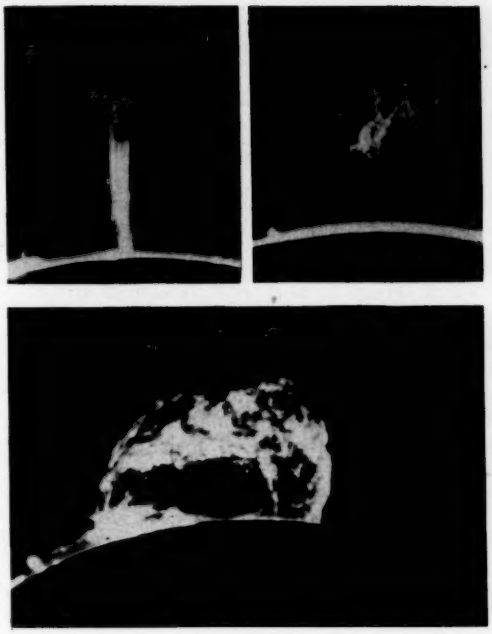

Plate I. Photographs of prominences at the limb of the sun taken through the birefringent filter of the High Altitude Observatory in the light of the $H\alpha$-line of hydrogen.

Lyot[3] has calculated the off-axis effect for light incident in the two principal planes normal to the α- and γ-axes in a biaxial crystal cut with its surfaces normal to the β-axis. Although the equations are not exact, since terms of the fourth and higher degrees in φ (the angle of incidence) are neglected, the approximation is excellent for the moderate angles of incidence encountered in the use of filters.

Lyot's equations can be very simply generalized to give the off axis effects for light incident in any plane normal to the surface of the crystal (and parallel, therefore, to the β-axis). Figure 3 represents a block of biaxial crystal with its α-, β-, and γ-axes in the directions indicated. Let polarized light with vibrations in a plane at 45° to the α-axis enter the crystal in the direction (φ, θ). Here φ is the angle of incidence, and θ is the azimuth of the incident plane measured from the α-axis. The light emerges from the crystal in the direction (φ, θ) in two polarized components with vibrations very closely parallel to the α- and γ-axes. They have a relative retardation, n, which is to be determined as a function of φ, θ, and n_0, where n_0 is the retardation for light entering the crystal from the direction $\varphi=0$.

A consideration of the isochromatic surfaces of

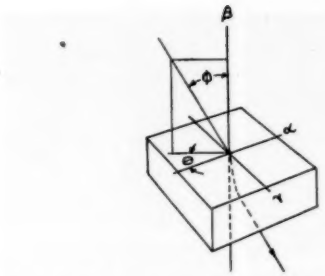

FIG. 3. Off-axis ray in the crystal coordinate system.

biaxial crystals[4] leads to the conclusion that the equations of the curves of constant retardation, n (written in terms of φ and θ), represent hyperbolas if terms in the fourth and higher powers of φ are neglected. Their transverse axes are along the α-axis for $n/n_0 \gtrless 1$ and along the γ-axis for $n/n_0 \lessgtr 1$ for crystals in which $\alpha\gamma-\beta^2 \gtrless 1$. The asymptotes are the lines

$$\tan^2\theta = \alpha/\gamma. \qquad \text{(III.1)}$$

Lyot's equations give the squares of the semi-transverse axes, which are:

$$\varphi_0^2 = \left(\frac{n}{n_0}-1\right)\frac{\gamma}{k} \quad \text{in the plane} \quad \theta=0,$$
$$\varphi_{\pi/2}^2 = \left(\frac{n}{n_0}-1\right)\frac{\alpha}{k} \quad \text{in the plane} \quad \theta=\pi/2, \qquad \text{(III.2)}$$

where

$$k = \frac{\alpha\gamma-\beta^2}{2(\gamma-\alpha)\beta^2}. \qquad \text{(III.3)}$$

We have, therefore, sufficient information to determine both sets of hyperbolas, which can be represented by a single equation

$$n = n_0\left[1+\varphi^2 k\left(\frac{\cos^2\theta}{\gamma}-\frac{\sin^2\theta}{\alpha}\right)\right]. \qquad \text{(III.4)}$$

The exact expression for n in uniaxial crystals is readily derived by a straightforward application of Huygens' principle and analytic geometry. Consider a plane-parallel uniaxial crystal in a rectangular x, y, z coordinate system with the origin in the first surface. Let it be oriented with its surfaces normal to the z axis. Let the x axis be parallel to the crystal optic axis (i.e., parallel to the α-axis in negative crystals or to the γ-axis in positive crystals). Choose units of time and distance to make the velocity of light in space unity. The equation of an entering plane light wave is then:

$$ax+by+cz-t=0, \qquad \text{(III.5)}$$

where a, b, and c are the direction cosines of the normal to the wave front and t is the time.

As the wave passes the origin in entering the crystal, it initiates a secondary wavelet which expands into an ellipsoid with the equation:

$$\zeta^2x^2+\eta^2y^2+v^2z^2-t^2=0, \qquad \text{(III.6)}$$

where ζ, η, and v are reciprocals of the velocities along the x, y, and z directions, respectively.

At a given instant, that portion of the plane wave which is inside the crystal, coincides with a plane tangent to the ellipsoid of Eq. (III.6) and containing the line of intersection of the plane wave of Eq. (III.5) with the first surface of the crystal (i.e., the plane $z=0$). The tangent plane through the point (x_1, y_1, z_1) on the ellipsoid is

$$x_1\zeta^2x+y_1\eta^2y+z_1v^2z-t^2=0. \qquad \text{(III.7)}$$

The lines of intersection of the planes of Eqs. (III.5) and (III.7) with the first surface of the crystal are, respectively,

$$ax+by-t=0, \quad z=0, \qquad \text{(III.8)}$$

and

$$x_1\zeta^2x+y_1\eta^2y-t^2=0, \quad z=0. \qquad \text{(III.9)}$$

These two lines must coincide. Hence:

$$x_1=(a/\zeta^2)t \quad \text{and} \quad y_1=(b/\eta^2)t. \qquad \text{(III.10)}$$

Since (x_1, y_1, z_1) must be a point on the ellipsoid of Eq. (III.6), we find for z_1:

$$z_1=\frac{t}{v}\left(1-\frac{a^2}{\zeta^2}-\frac{b^2}{\eta^2}\right)^{\frac{1}{2}}. \qquad \text{(III.11)}$$

Equations (III.10) and (III.11) define the path of a ray through the origin.

Let d be the thickness of the crystal in the z direction. The time, t_1, when a ray through the origin reaches the second surface is, then:

$$t_1=dv/[1-(a^2/\zeta^2)-(b^2/\eta^2)]^{\frac{1}{2}}. \qquad \text{(III.12)}$$

On emerging from the crystal the plane wave is parallel to the entering wave, with the equation:

$$ax+by+cz-(t-\Delta)=0. \qquad \text{(III.13)}$$

At time, t_1, this plane must contain the point (x_1, y_1, d). Hence,

$$\Delta=t_1-(ax_1+by_1+cd).$$

The distance, p, of the plane wave of Eq. (III.13) from the origin is therefore:

$$p=t-\Delta=t-t_1+ax_1+by_1+cd, \qquad \text{(III.14)}$$

or, from Eqs. (III.10) and (III.12):

$$p=t-d\left[v\left(1-\frac{a^2}{\zeta^2}-\frac{b^2}{\eta^2}\right)^{\frac{1}{2}}-c\right]. \qquad \text{(III.15)}$$

[4] T. Preston, *Theory of Light*, 3rd Edition (MacMillan Company, Ltd., London, 1901), p. 397.

Now, for the extraordinary wave

$$\zeta=\omega, \quad \eta=\upsilon=\epsilon,$$

and for the ordinary wave

$$\zeta=\eta=\upsilon=\omega.$$

Hence, the distances of the extraordinary and ordinary waves from the origin after their traversal of the crystal can be written, respectively:

$$\begin{aligned} p_\epsilon &= t-d\left[\epsilon\left(1-\frac{a^2}{\omega^2}-\frac{b^2}{\epsilon^2}\right)^{\frac{1}{2}}-c\right], \\ p_\omega &= t-d\left[\omega\left(1-\frac{a^2+b^2}{\omega^2}\right)^{\frac{1}{2}}-c\right]. \end{aligned} \tag{III.16}$$

The retardation, n, is simply $(p_\omega-p_\epsilon)/\lambda$, or:

$$n=\frac{n_0}{\epsilon-\omega}\left[\epsilon\left(1-\frac{a^2}{\omega^2}-\frac{b^2}{\epsilon^2}\right)^{\frac{1}{2}}-\omega\left(1-\frac{a^2+b^2}{\omega^2}\right)^{\frac{1}{2}}\right]. \tag{III.17}$$

Equation (III.17) is the exact expression for the off-axis effect in uniaxial crystals. It is readily reduced to the more convenient approximation of Eq. (III.4). Expanding the radicals, and neglecting fourth and higher powers of a and b, we find:

$$n=\frac{n_0}{\epsilon-\omega}\left[\epsilon-\omega-\frac{\epsilon}{2\omega^2}a^2-\frac{1}{2\epsilon}b^2+\frac{1}{2\omega}(a^2+b^2)\right]. \tag{III.18}$$

The direction cosines can be expressed in terms of φ and θ by the transformation:

$$a=\sin\varphi\sin\theta'; \quad b=\sin\varphi\cos\theta',$$

where

$$\theta'=\theta \quad \text{if} \quad \epsilon-\omega>0,$$

and

$$\theta'=\theta+\pi/2 \quad \text{if} \quad \epsilon-\omega<0.$$

Equation (III.18) becomes, then:

$$n=n_0\left[1+\frac{\varphi^2}{2\omega}\left(\frac{\cos^2\theta'}{\epsilon}-\frac{\sin^2\theta'}{\omega}\right)\right]. \tag{III.19}$$

Equation (III.19) is identical with Eq. (III.4) for uniaxial crystals.

The corresponding exact equation for biaxial crystals can be derived by the same methods, but the resulting expressions become so lengthy and complicated that it has not seemed worth while to push them through. The accuracy of Eq. (III.4) is adequate for all practical purposes whether uniaxial or biaxial crystals are considered.

It should be noted that the use of the equations of isochromatic surfaces in the derivation of off-axis effects does not lead to an exact result, since they are derived on the inexact assumption that the two components of light polarized at right angles traverse the crystal along identical paths.

The usable field of a given filter is determined by the maximum permissible value of $|n-n_0|$ for the thickest b-element.

The maximum permissible angle of incidence in the Climax filter in the $\theta=\pi/2$ plane is

$$\varphi=0.025 \text{ radian},$$

if we require that over the field

$$|n-n_0|\leqslant 0.1$$

for the thickest b-element.

IV. LYOT'S WIDE FIELD FILTERS

The maximum total flux from a given light source that can be squeezed through a filter is roughly proportional to the square of the product of the filter aperture and the maximum usable value of φ. The aperture is limited by the sizes of available birefringent crystals, and it is therefore important to find means for obtaining large fields. The most obvious device is to find a birefringent material for which k is very small. Although the author knows of no such material which is available in useful sizes of optical quality, this is a definite possibility which should be investigated further.

Lyot[3] has described three wide-field filters with compound elements made of available materials. They will be referred to as Lyot's first-type, second-type, and third-type filters.

The first-type filter differs from the simple filter in having each b-element divided into two equal halves by a cut perpendicular to the β-axis. The second half of each element is rotated about the β-axis until the α-axes of the two components are crossed. A half-wave plate is inserted between the components with its α-axis at 45° to the α-axes of the two. It serves to rotate the planes of polarization 90°. Light which enters the first component from the direction (φ, θ) enters the second component from the direction $(\varphi, \theta+\pi/2)$. The retardation introduced by the assembled element is then:

$$\begin{aligned} n &= \tfrac{1}{2}\left[n(\varphi,\theta)+n\left(\varphi,\theta+\frac{\pi}{2}\right)\right] \\ &= \tfrac{1}{2}n_0\left[1+\varphi^2k\left(\frac{\cos^2\theta}{\gamma}-\frac{\sin^2\theta}{\alpha}\right)\right] \\ &\quad +\tfrac{1}{2}n_0\left[1+\varphi^2k\left(\frac{\sin^2\theta}{\gamma}-\frac{\cos^2\theta}{\alpha}\right)\right], \end{aligned}$$

or

$$n=n_0\left[1+\varphi^2\frac{k}{2}\left(\frac{1}{\gamma}-\frac{1}{\alpha}\right)\right]. \tag{IV.1}$$

The loci of constant retardation are circles with radii larger than the axes of the hyperbolas of a simple filter (in the $\theta=\pi/2$ plane) by a factor of

$[2\gamma/(\gamma-\alpha)]^{\frac{1}{2}}$. For a given tolerable value of $|n-n_0|$, the radius of the useful field can be further increased by a factor of $\sqrt{2}$ if we set the retardation at the center of the field at one extreme of the range.

Lyot's first-type filter, unlike the simple filter, can be used only over a small range of wave-lengths. If the wave-length differs greatly from the optimum for which the half-wave plates are made, the residual light between transmission bands increases at the expense of light in the bands. The added residual light appears superposed on the field in the form of faint hyperbolic fringes very similar to the fringes produced by the equivalent simple filter. The fringes are loci of constant retardation, n', given by

$$n'=\frac{n_0}{2}k\varphi^2\left(\frac{1}{\gamma}+\frac{1}{\alpha}\right)\cos 2\theta. \qquad \text{(IV.2)}$$

If, however, the filter is either used for one wave-length only, or supplied with interchangeable half-wave plates for the different spectral regions, its performance is highly satisfactory. This is one of the many instances where the development of an achromatic half-wave plate would be very useful.

Lyot's second-type wide-field filter has compound b-elements of two components of different materials. The quantity k is of opposite sign in the two components, which are mounted with their α-axes parallel. No half-wave plates are required.

Let n_1 and n_2 be the retardations arising from the first and second components for light entering from the direction $\varphi=0$. The retardation for the assembled element is then:

$$n=n_1\left[1+\varphi^2k_1\left(\frac{\cos^2\theta}{\gamma_1}-\frac{\sin^2\theta}{\alpha_1}\right)\right]+n_2\left[1+\varphi^2k_2\left(\frac{\cos^2\theta}{\gamma_2}-\frac{\sin^2\theta}{\alpha_2}\right)\right],$$

$$n=n_0+\varphi^2\left[\cos^2\theta\left(\frac{n_1k_1}{\gamma_1}+\frac{n_2k_2}{\gamma_2}\right)-\sin^2\theta\left(\frac{n_1k_1}{\alpha_1}+\frac{n_2k_2}{\alpha_2}\right)\right], \qquad \text{(IV.3)}$$

where now

$$n_0=n_1+n_2. \qquad \text{(IV.4)}$$

It is evident that while the coefficient of φ^2 cannot be made to vanish by any choice of n_1 and n_2, we can obtain circular fringes by eliminating θ. The condition is

$$\frac{n_1}{n_2}=-\left(\frac{k_2}{k_1}\right)\left[\left(\frac{1}{\gamma_2}+\frac{1}{\alpha_2}\right)\Big/\left(\frac{1}{\gamma_1}+\frac{1}{\alpha_1}\right)\right]. \qquad \text{(IV.5)}$$

Equations (IV.4) and (IV.5) give n_1 and n_2. The retardation of the assembled element can now be written

$$n=n_0+\varphi^2\left(\frac{n_1k_1}{\gamma_1}+\frac{n_2k_2}{\gamma_2}\right). \qquad \text{(IV.6)}$$

The second-type filter can be used over a wide range of wave-lengths, although the fringes do not remain strictly circular throughout the range.

Lyot's third-type filter generally has the largest field. Each b-element consists of three birefringent components. Two of the components are of the same material and are mounted with their α-axes crossed. The third is of a different birefringent material with a k value opposite in sign to the k value for the first two components. It is mounted with its α-axis parallel to that of one of the first two. By a proper choice of thicknesses it is always possible to make n constant over the whole field within the accuracy of Eq. (III.4).

Let α_1, β_1, γ_1 and α_2, β_2, γ_2 be the refractive indices of the crystals composing the single component and the two crossed components, respectively. The crystals must be selected to satisfy the condition

$$\alpha_1\gamma_2>\gamma_1\alpha_2.$$

Let n_a be the retardation of the single component, and n_b and n_c the retardations of the two components of the same material. Let the α-axes of the a and b components be in the $\theta=0$ plane, and the α-axis of the c component in the $\theta=\pi/2$ plane. Then

$$n=n_a\left[1+\varphi^2k_1\left(\frac{\cos^2\theta}{\gamma_1}-\frac{\sin^2\theta}{\alpha_1}\right)\right]+n_b\left[1+\varphi^2k_2\left(\frac{\cos^2\theta}{\gamma_2}-\frac{\sin^2\theta}{\alpha_2}\right)\right]-n_c\left[1+\varphi^2k_2\left(\frac{\sin^2\theta}{\gamma_2}-\frac{\cos^2\theta}{\alpha_2}\right)\right]. \qquad \text{(IV.7)}$$

If we set $n_a+n_b-n_c=n_0$, and require that the coefficient of φ^2 vanish, we find

$$n_a=\frac{n_0}{A}k_2^2\left(\frac{1}{\alpha_2^2}-\frac{1}{\gamma_2^2}\right),$$

$$n_b=\frac{n_0}{A}k_1k_2\left(\frac{1}{\gamma_1\gamma_2}-\frac{1}{\alpha_1\alpha_2}\right), \qquad \text{(IV.8)}$$

$$n_c=\frac{n_0}{A}k_1k_2\left(\frac{1}{\alpha_1\gamma_2}-\frac{1}{\gamma_1\alpha_2}\right),$$

where

$$A=\begin{vmatrix} \frac{k_1}{\gamma_1} & \frac{k_2}{\alpha_2} & \frac{k_2}{\gamma_2} \\ \frac{k_1}{\alpha_1} & \frac{k_2}{\gamma_2} & \frac{k_2}{\alpha_2} \\ 1 & -1 & 1 \end{vmatrix}. \qquad \text{(IV.9)}$$

The retardation of the assembled element for any direction (φ, θ) is then:

$$n=n_a+n_b-n_c=n_0. \qquad \text{(IV.10)}$$

The third-type filter, like the second-type, can be used over a wide range of wave-lengths. The coefficient of φ^2, however, will generally vanish accurately at only one wave-length.

The design of a wide angle filter does not necessarily require that the thinner elements be compound. Their transmission bands are so broad that the slight shift in wave-length for off-axis rays is negligible in comparison. If the higher order compound elements are made of two materials, however, it may not be possible to use transmission bands in widely separated regions of the spectrum, because the dispersions of different materials are generally not strictly proportional. If the rth element is the thickest simple element, $(n_{r+1})/n_r=2$ at only one wave-length.

The following sections are devoted to the theory of various modifications of birefringent filters which have been recently developed.

V. THE SPLIT ELEMENT FILTER

The split-element filter resembles Lyot's first-type filter, and shares its wide field characteristics. The half-wave plates, however, are replaced by birefringent elements, and successive polarizers are crossed. After the initial polarization, it requires only half as many polarizers as the equivalent simple filter. The result is a considerable reduction in absorption and scattered light if film polarizers are used, or a notable saving in bulk and expense if polarizing prisms are used.

The split-element filter has already been described briefly.[5] A more detailed account of its theory is given here.

A single unit of the split-element filter (which would be mounted between crossed polarizers) is shown schematically in Fig. 4. The x, y, and z axes constitute a rectangular coordinate system. The positive r and s axes in the xy plane bisect the angles between the positive x and y and the positive y and negative x directions, respectively. The unit

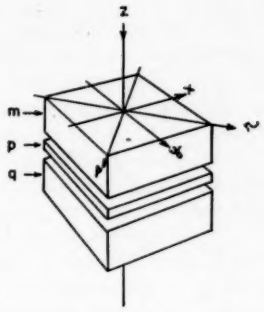

FIG. 4. Birefringent components of a single unit of a split element filter.

consists of a split element with components m and q, and a simple element, p, sandwiched between m and q. They are all mounted with β-axes parallel to the z axis. The γ-axes are aligned parallel to the x, r, and y directions, respectively, in the m, p, and q components. Let the thicknesses of m, p, and q be d_m, d_p, and d_q, and let the unit of time be the vibration period of the light.

Assume that the entering light is polarized in the r plane. The transmissions of the unit for emerging light polarized in the r plane and s plane are to be determined.

The vibration of the entering light is

$$r=a\sin 2\pi t. \qquad \text{(V.1)}$$

This can be resolved along the x and y directions giving

$$x=(a/\sqrt{2})\sin 2\pi t, \quad y=(a/\sqrt{2})\sin 2\pi t. \qquad \text{(V.2)}$$

In traversing m, a phase difference is introduced and the vibration of the emerging light is

$$\begin{aligned} x_m&=(a/\sqrt{2})\sin 2\pi(t-d_m\gamma), \\ y_m&=(a/\sqrt{2})\sin 2\pi(t-d_m\alpha). \end{aligned} \qquad \text{(V.3)}$$

The resultant disturbance along the r and s axes is:

$$\begin{aligned} r_m&=a\cos\pi n_m\sin 2\pi t', \\ s_m&=a\sin\pi n_m\cos 2\pi t', \end{aligned} \qquad \text{(V.4)}$$

where

$$t'=t-(d_m/2\lambda)(\alpha+\gamma).$$

In the traversal of p, an additional phase difference is introduced;

$$\begin{aligned} r_p&=a\cos\pi n_m\sin 2\pi[t'-(d_p/\lambda)\gamma], \\ s_p&=a\sin\pi n_m\cos 2\pi[t'-(d_p/\lambda)\alpha]. \end{aligned} \qquad \text{(V.5)}$$

Resolving this vibration along the x and y axes, and adding the phase difference due to transmission

[5] J. Evans, Ciencia e Investigación (Argentina) Vol. III, No. 9, p. 365 (1947).

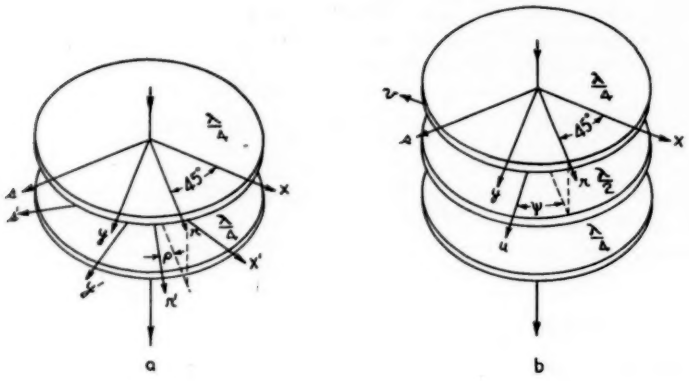

FIG. 5. (a) Phase shifter of two quarter-wave plates. (b) Phase shifter of one half-wave and two quarter-wave plates.

through q, we obtain

$$x_q=\frac{a}{\sqrt{2}}\cos\pi n_m \sin 2\pi\left(t'-\frac{d_p}{\lambda}\gamma-\frac{d_q}{\lambda}\alpha\right)$$
$$-\frac{a}{\sqrt{2}}\sin\pi n_m \cos 2\pi\left(t'-\frac{d_p}{\lambda}\alpha-\frac{d_q}{\lambda}\alpha\right), \tag{V.6}$$
$$y_q=\frac{a}{\sqrt{2}}\cos\pi n_m \sin 2\pi\left(t'-\frac{d_p}{\lambda}\gamma-\frac{d_q}{\lambda}\gamma\right)$$
$$+\frac{a}{\sqrt{2}}\sin\pi n_m \cos 2\pi\left(t'-\frac{d_p}{\lambda}\alpha-\frac{d_q}{\lambda}\gamma\right).$$

To determine the final transmission through a polarizer with its axis along either the r or s direction, we must resolve this vibration along the r and s axes:

$$r_q=a\cos\pi n_m\cos\pi n_q\sin 2\pi[t''-(d_p/\lambda)\gamma]$$
$$+a\sin\pi n_m\sin\pi n_q\sin 2\pi[t''-(d_p/\lambda)\alpha], \tag{V.7}$$
$$s_q=a\sin\pi n_m\cos\pi n_q\cos 2\pi[t''-(d_p/\lambda)\gamma]$$
$$-a\cos\pi n_m\sin\pi n_q\cos 2\pi[t''-(d_p/\lambda)\alpha],$$

where

$$t''=t-[(d_m+d_q)/2\lambda](\alpha+\gamma).$$

Let the emergent amplitudes be A_r and A_s. The transmissions in the r and s vibration planes are, then:

$$\tau_r=A_r^2/a^2=\cos^2\pi(n_m-n_q)$$
$$-\sin 2\pi n_m\sin 2\pi n_q\sin^2\pi n_p, \tag{V.8}$$
$$\tau_s=A_s^2/a^2=\sin^2\pi(n_m-n_q)$$
$$+\sin 2\pi n_m\sin 2\pi n_q\sin^2\pi n_p.$$

In the split-element filter the m and q components are made of equal thicknesses. Hence

$$n_m=n_q.$$

If we let

$$n_j=2n_m=2n_q,$$

Eqs. (V.8) reduce to:

$$\tau_r=1-\sin^2\pi n_j\sin^2\pi n_p, \tag{V.9}$$
$$\tau_s=\sin^2\pi n_j\sin^2\pi n_p.$$

The transmission of an element of Lyot's first-type filter is τ_r in Eqs. (V.9) if we set $n_p=\frac{1}{2}$.

The transmission of a unit of the split-element filter is τ_s. A split-element filter of l elements has exactly the same off-axis characteristics as a filter of Lyot's first type with the first $l/2$ elements simple, and the $l/2$ thicker elements compound. Whether the field is limited by the simple elements or the compound elements depends upon whether or not $(n_l/n_{l/2})[(\gamma-\alpha)/(2\gamma)]$ is greater or less than 1. If the simple elements limit the field, they can, of course, be made compound in any of Lyot's three types.

The transmission of an assembled split-element filter composed of two-element units between crossed polarizers is:

$$\tau=\sin^2\pi n_1\sin^2\pi n_2\cdots\sin^2\pi n_l. \tag{V.10}$$

Since transmission bands occur only at wave-lengths for which all the n's are half-integral, the n's cannot be simply proportional to the powers of 2. If we let $n=n'+\frac{1}{2}$ at the wave-length of a particular band, the best we can do is to make the values of n' proportional to the powers of 2. Thus

$$n_r=2^{r-1}n_1'+\tfrac{1}{2}. \tag{V.11}$$

The transmission can then be written

$$\tau=\cos^2\pi n_1'\cos^2\pi 2n_1'\cdots\cos^2\pi 2^{l-1}n_1'. \tag{V.12}$$

Unfortunately Eq. (V.11) can be strictly valid at only one wave-length, and the usefulness of the filter is restricted to a limited spectral region in the neighborhood of that wave-length. This is a second instance where achromatic half wave plates would be useful. If the r-th element of the filter were made to give a retardation $n_r'=2^{r-1}n_1'$, the addition of an achromatic half-wave plate (two quarter-wave plates for split elements) would satisfy Eq. (V.11) at all wave-lengths.

The thought will doubtless have occurred to the reader that the middle element in each unit of a split-element filter could itself be split, and a third

element inserted between the halves. This plan does not work theoretically, and so far no arrangement has been found that allows more than two elements in a unit between successive polarizers.

VI. FILTERS OF ADJUSTABLE WAVE-LENGTH

It is obvious that the usefulness of the birefringent filter is enormously enhanced if a transmission maximum can be adjusted to center on any desired wave-length. The fine adjustment resulting from the control of temperature is generally quite inadequate as it has a range of only a few angstroms (although Lyot found that with the aid of temperature control it is possible to bring no less than six of the maxima of a quartz filter into coincidence with lines of major importance in the solar spectrum).

The obvious method of controlling the wave-length of the transmission bands is by means of elements of variable thickness, made of pairs of wedges which can be adjusted with respect to each other like the components of a Babinet compensator. It is then possible to set

$$n_1 = \text{an integer},$$

and

$$n_r = 2^{r-1} n_1$$

at any chosen wave-length. Such an arrangement is perfectly feasible and works equally well at all wave-lengths. In the split-element filter, both halves of the split element must, of course, be adjustable since $n_m - n_q = 0$. The range of variation in thickness need be only sufficient to shift the principal transmission maxima of the filter through a range equal to their separation. With a proper choice of wedge angles all the movable wedges can be mounted and adjusted as a single unit.

Although theoretically excellent, the variable-thickness filter requires considerable mechanical refinement, and one wedge in each element must have an aperture much larger than the instrumental aperture (a matter of importance in filters of large aperture). The use of phase shifters for wave-length adjustment is simpler and, for most purposes, equally satisfactory. If achromatic phase shifters can be devised, they will give results as theoretically perfect as variable thickness.

Suppose we equip each b-element of a filter with a phase shifter which permits the addition of a small controllable phase difference, $2\pi\xi$, to the phase difference, $2\pi n$, introduced by the b-element. The transmission of the filter is then

$$\tau = \prod_{r=1}^{r=l} \cos^2 \pi (n_r + \xi_r). \qquad \text{(VI.1)}$$

Again, with the split-element filter, the added phase difference must be divided equally between the two halves of the split elements to keep $n_m - n_q = 0$. A transmission maximum of the filter can then be centered on any given wave-length, λ_1, by adjusting ξ until $n + \xi$ is an integer for each element. This is always possible if ξ can be adjusted over the range $-\frac{1}{2}$ to $+\frac{1}{2}$. If the phase shifter is achromatic, i.e., ξ is independent of wave-length at a given setting, the result is merely a shift of the transmission curve of the filter along the spectrum and its performance is equally good at all wave-length settings. If, on the other hand, ξ is a function of wave-length, the spacings of the transmission maxima of a given element are altered. Hence the relative positions of the transmission maxima and minima of the different elements depart more and more from exact superposition as the wave-length departs from λ_1. The result is an increase in the residual light transmitted in the intervals between principal maxima of the filter as $|\lambda - \lambda_1|$ increases.

Lyot[3] and Billings[6] have both made numerical calculations of the additional residual light resulting from the use of non-achromatic phase shifters. They concluded that over a reasonable wave-length range (which can readily be isolated with glass or gelatine filters) the increase in residual light is negligible. The adjustment of wave-length with phase shifters is therefore a practical possibility whether the phase shifters are achromatic or not.

Several forms of variable phase shifters have been proposed.

Lyot[3] made elements of variable thickness like

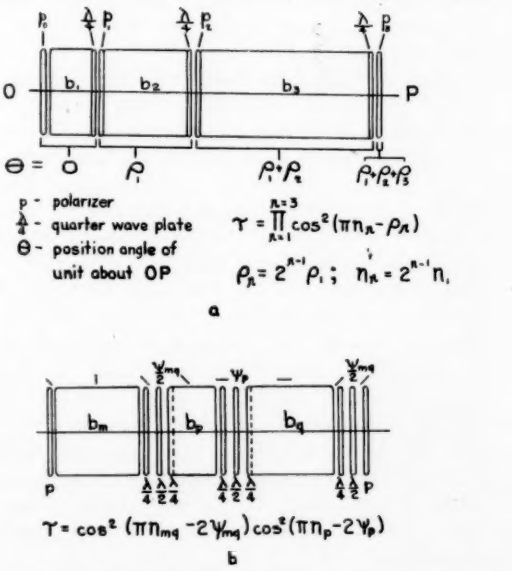

FIG. 6. (a) Simple filter of three elements with quarter-wave plate phase shifters. (b) One unit of a split element filter with fractional wave plate phase shifters.

[6] Bruce H. Billings, J. Opt. Soc. Am. **37**, 738 (1947).

those described above for the variable thickness filter, but with the difference that the range of adjustment of retardation was restricted to one wave-length.

Billings[6] made an experimental filter with photoelastic phase shifters composed of sheets of polyvinyl butyrate under adjustable tension.

While both these arrangements give a satisfactory wave-length adjustment, they are tedious to use. Ordinarily each element must be individually adjusted. The alternative is a complicated mechanical synchronization of the adjustments of all the elements, which would make operation with a single control feasible. Without some such arrangement it would be impossible to vary the wave-length continuously.

A much more promising approach is the use of the electro-optical phase shifters discussed by Billings.[6] A plate of the uniaxial crystal ammonium di-hydrogen phosphate ($NH_4H_2PO_4$), known commercially as PN, cut perpendicular to the optic axis and mounted between transparent electrodes, becomes biaxial and exhibits a retardation when a potential difference is applied to the electrodes. The retardation is proportional to the potential difference and is independent of the thickness of the PN plate. A filter made with a Billings plate added to each element (to each half of the split elements in the split element filter) could be adjusted electrically, and the problem of synchronizing the phase shifts of successive elements would be relatively simple. At the present writing Dr. Billings is actively engaged in the development of such electrically tunable filters.

All three tuning methods have one difficulty in common. It is impracticable to push the phase shift beyond a very limited range. If a range from $-\pi$ to $+\pi$ is adopted, a continuous variation of wave-length involves a discontinuous adjustment of each phase shifter. The phase shift must progress smoothly from $-\pi$ to $+\pi$ (at a rate proportional to the thickness of the associated b-element) and then jump back to $-\pi$. For most purposes there may be no serious disadvantage in this. If, however, the filter is to be used for spectrophotometric work, for example, it may be very difficult to avoid a spurious bump in the filter transmission every time a phase shifter passes a point of discontinuity, even with the electrical tuning. For such special purposes phase shifters composed of rotating fractional wave plates can be used. They have already been described briefly.[5] A fuller account of their theory is given here.

The specific problem is to devise a combination of fractional wave plates which will alter the phase difference between the vibrations along two mutually perpendicular axes, x and y, by any chosen amount, without altering their amplitudes. At a given wave-length, this is equivalent to a variable thickness of birefringent material with its γ-axis along the x or y direction. Such an arrangement is shown at a, Fig. 5. It consists of two quarter-wave plates. The first is fixed with its γ-axis along the r axis (at 45° to the x axis). The second can be rotated around the instrumental axis. At a given setting its γ-axis lies along the r' direction at angle ρ to the r direction.

The vibration of the light entering the system is generally represented by

$$x = b \sin 2\pi t, \quad y = c \sin 2\pi(t+\sigma). \tag{VI.2}$$

Resolving this vibration along the r and s axes and adding a phase difference of $\pi/2$ introduced by the first quarter-wave plate we obtain for the emerging vibration:

$$\begin{aligned} r &= (b/\sqrt{2}) \sin 2\pi t + (c/\sqrt{2}) \sin 2\pi(t+\sigma), \\ s &= -(b/\sqrt{2}) \cos 2\pi t + (c/\sqrt{2}) \cos 2\pi(t+\sigma). \end{aligned} \tag{VI.3}$$

Resolving this vibration along the r' and s' axes and adding another phase difference of $\pi/2$ introduced by the second quarter-wave plate, we obtain,

$$\begin{aligned} r' &= (b/\sqrt{2}) \sin[2\pi t - \rho] \\ &\quad + (c/\sqrt{2}) \sin[2\pi(t+\sigma) + \rho], \\ s' &= (b/\sqrt{2}) \sin[2\pi t - \rho] \\ &\quad - (c/\sqrt{2}) \sin[2\pi(t+\sigma) + \rho]. \end{aligned} \tag{VI.4}$$

Finally, if we resolve this vibration along the x' and y' axes, at an angle of $\rho + (\pi/2)$ to the x and y axes, we obtain for the emerging vibration:

$$\begin{aligned} x' &= b \sin[2\pi t - \rho], \\ y' &= c \sin[2\pi(t+\sigma) + \rho + \pi]. \end{aligned} \tag{VI.5}$$

A comparison of Eqs. (VI.2) with (VI.5) shows that while the emerging amplitudes along x' and y' are the same as the entering amplitudes along x and y, the phase difference has been increased from $2\pi\sigma$ to $2\pi\sigma + 2\rho + \pi$, i.e., the phase shift, $2\pi\xi$, is

$$2\pi\xi = \pi + 2\rho. \tag{VI.6}$$

Obviously the phase difference can be set to any desired value by adjusting ρ.

This two-element phase shifter has the disadvantage that the x' and y' axes rotate with the second quarter-wave plate. For some applications this is no inconvenience but in others it renders this phase shifter useless. The x' and y' axes can be restored to parallelism with the x and y axes by the addition of a rotatable half-wave plate, which has the property of reflecting any polarization figure in its γ-axis.

The most convenient system, shown at b, Fig. 5, consists of two fixed quarter-wave plates with the rotatable half-wave plate sandwiched between them. Suppose the γ-axes of both quarter-wave plates are in the r direction, while the γ-axis of the

half-wave plate is along the u direction at an angle ψ to the r direction.

The vibration emerging from the first quarter-wave plate is given by Eq. (VI.3). Resolving this vibration along the u and v axes, and adding a phase difference of π, we obtain for the vibration emerging from the half-wave plate:

$$\begin{aligned} u&=(b/\sqrt{2})\sin[2\pi t-\psi] \\ &\quad+(c/\sqrt{2})\sin[2\pi(t+\sigma)+\psi], \\ v&=(b/\sqrt{2})\cos[2\pi t-\psi] \\ &\quad-(c/\sqrt{2})\cos[2\pi(t+\sigma)+\psi]. \end{aligned} \tag{VI.7}$$

Resolving this vibration again along the r and s axes, and adding a phase difference of $\pi/2$, we obtain for the vibration emerging from the second quarter-wave plate:

$$\begin{aligned} r&=(b/\sqrt{2})\sin[2\pi t-2\psi] \\ &\quad+(c/\sqrt{2})\sin[2\pi(t+\sigma)+2\psi], \\ s&=-(b/\sqrt{2})\sin[2\pi t-2\psi] \\ &\quad+(c/\sqrt{2})\sin[2\pi(t+\sigma)+2\psi]. \end{aligned} \tag{VI.8}$$

Finally, resolving this vibration along the original x and y axes, we find

$$\begin{aligned} x&=b\sin[2\pi t-2\psi], \\ y&=c\sin[2\pi(t+\sigma)+2\psi]. \end{aligned} \tag{VI.9}$$

The phase shift introduced by the three-element system is, therefore,

$$2\pi\xi=4\psi. \tag{VI.10}$$

The principal advantage in the use of fractional wave plate phase shifters in birefringent filters is in the possibility of a continuous variation of wave-length without discontinuities in the adjustment of the moving elements. Since ρ or ψ can be increased or decreased indefinitely, $2\pi\xi$ is not restricted as it is in the other types of phase shifters discussed above.

It should be noted that the fractional wave plate phase shifter is in a sense achromatic, since ξ is independent of the wave-length for a given value of ρ or ψ—a very desirable property (see the discussion following Eq. (VI.1)). With ordinary quarter- and half-wave plates, however, this advantage is somewhat illusory. Their usefulness is limited to the rather restricted region of the spectrum where their retardations are very close to quarter-wave and half-wave. This is another application where the desirability of achromatic fractional wave plates is evident.

If continuity of adjustment over a large range of the spectrum is a necessity, the fractional wave plates themselves could be made adjustable. The addition of an electro-optical Billings plate to each fractional wave plate would perhaps be the simplest method. A relatively moderate potential applied to the Billings plate would then adjust the retardation accurately to a half-wave or quarter-wave at the wave-length of the transmission band of the filter. This seems a rather desperate measure, however.

The construction of the fractional wave plate phase shifters is considerably simplified when they are used in birefringent filters. Some of the quarter-wave plates simply take the form of an addition to the thickness of the birefringent elements. In instances where the γ-axis of a quarter-wave plate is parallel or perpendicular to the axis of an immediately following polarizer, it is evident that the polarizer utilizes only one component of the vibration emerging from the quarter-wave plate. The $\pi/2$ phase difference therefore serves no real purpose, and the quarter-wave plate can be omitted.

Consider first an element of a simple filter. Suppose the b-element, oriented with its γ-axis along the x direction, is followed by a quarter-wave plate with its γ-axis along the r direction. If we let $b=c=a/\sqrt{2}$, $t=t'-(d/2\lambda)\mu$, and $\sigma=(d/\lambda)\mu$, Eq. (VI.3) for the vibration emerging from the quarter-wave plate reduces to:

$$r=a\cos\pi n\sin 2\pi t', \quad s=a\sin\pi n\sin 2\pi t'. \tag{VI.11}$$

This is a linear vibration at an angle of πn to the r-axis. We can omit the second quarter-wave plate and let the light enter a polarizer with its plane of polarization at angle ρ to the r axis. The transmission of the assembly is then

$$\tau=\cos^2(\pi n-\rho). \tag{VI.12}$$

By adjusting ρ (i.e., by rotating the polarizer) until $n_\lambda-\rho/\pi=$an integer, we can set $\tau=1$ for any chosen wave-length.

Lyot[3] has utilized this device to effect a slight shift in the wave-length of the transmission band of his filter. He used a quarter-wave plate with the last (thickest) element, and provided for the rotation of the final polarizer. The same method can be applied to the whole filter, however.

An adjustable simple birefringent filter would consist, then, of a series of units shown at a, Fig. 6, each composed of a polarizer, a birefringent element with its γ-axis at 45° to the axis of the polarizer, and a quarter-wave plate with its γ-axis parallel to the axis of the polarizer. The three parts of each unit remain fixed with respect to each other, but the unit itself must be rotatable around the instrumental axis. The angle ρ_r is then the angle between the γ-axis of the rth quarter-wave plate and the axis of the immediately following polarizer. The birefringent elements have the same thickness as in the non-adjustable filter. The transmission of the whole is

$$\begin{aligned} \tau=\cos^2(\pi n_1-\rho_1)\cos^2(\pi 2n_1-\rho_2)\cdots \\ \times\cos^2(\pi 2^{l-1}n_1-\rho_l), \end{aligned} \tag{VI.13}$$

and

$$\rho_2=2^{r-1}\rho_1. \tag{VI.14}$$

Since the values of ρ are proportional to the powers of 2, it is a relatively simple matter to devise a gear train by which the wave-length of the transmission band can be adjusted with a single control knob. A continuous variation of wave-length now involves no discontinuity in the adjustment of the various units, since ρ can be made to increase or decrease indefinitely.

Matters are somewhat more complicated in the split-element filter. The wide field characteristics depend upon the m and q components being crossed. Hence the phase shifts must be accomplished without any relative rotation of the two. Various arrangements are possible, some of which involve rotation of the center p-elements, or rotation of the unit as a whole with respect to the polarizers, or both. However, the unit shown at b, Fig. 6, is as simple as any.

The orientation of each element is indicated in the diagram by the short line above it for the fixed elements, or by the symbol $(\psi_{mq}/2)$ or ψ_p for the adjustable half-wave plates. The angle $(\psi_{mq}/2)$ or ψ_p is the angle between the γ-axis of the half-wave plate and the γ-axis of the preceding quarter-wave plate. The second quarter-wave plates following m and p are indicated as an addition to the thicknesses of p and q, while that following q has been omitted, since its γ-axis would be parallel to the axis of the following polarizer. The transmission of a split-element filter composed of such units is

$$\tau=\prod_{r=1}^{r=l}\cos^2\left(\pi n_r-2\psi_r-\frac{\pi}{2}\right). \qquad \text{(VI.15)}$$

It should be noted here that the built-in quarter-wave plates which are added to the thicknesses of the p- and q-elements are not included in the calculation of n for these elements.

The values of ψ_r should be proportional to n_r in Eq. (VI.15). Hence if $n_r=2^{r-1}(n_1-\frac{1}{2})+\frac{1}{2}$ as in the non-adjustable split-element filter, the ψ's are proportional to large odd numbers, and the problem of synchronizing the rotations of the half-wave plates becomes complicated (but not at all impossible). If, on the other hand, the n's are made proportional to the powers of two, the phase changers can compensate for the subtraction of $\frac{1}{2}$ from each value of n in addition to their normal function. Then

$$n_r=2^{r-1}n_1, \qquad \text{(VI.16)}$$

and

$$2\psi_r=\pi/2+2^{r-1}[2\psi_1-(\pi/2)]. \qquad \text{(VI.17)}$$

Since a rotation of the zero point from which angle ψ is measured to $\pi/4$ reduces this equation to

$$2\psi_r'=2^{r-1}(2\psi_1'), \qquad \text{(VI.18)}$$

it is evident that the variable parts of the ψ's are proportional to the powers of two, and the problem of synchronization becomes relatively simple.

The synchronization of the other types of phase shifters (variable thickness, photo elastic, or electro optical) is similarly simplified in a split element filter by constructing it with n's proportional to powers of two. Equations (VI.14) and (VI.17) apply if we substitute $\pi\xi$ for 2ψ.

A final remark about filters of adjustable wave-length seems worth while. The birefringent elements need not be made to any exact thicknesses as in the fixed wave-length filters. It is desirable, but not necessary, to preserve the relation $n_r=2^{r-1}n_1$ as closely as possible, since the synchronization of the various adjustments is then easier. There is no necessity, however, for n_1 to be an integer for any specified wave-length. This simplifies the construction somewhat. If $\mu\leqslant 0.03$, the thicknesses of the elements can be adjusted with sufficient accuracy by mechanical measurements alone. The error tolerance in thickness is inversely proportional to μ and is about ± 0.001 mm for $\mu=0.03$.

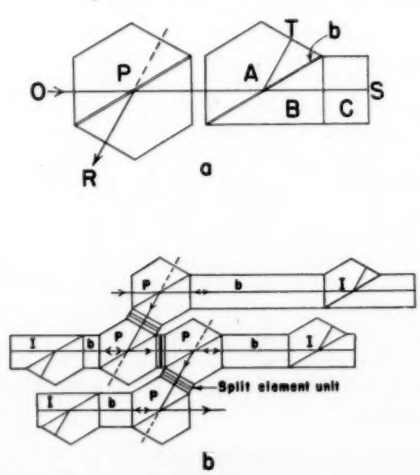

FIG. 7. (a) One form of polarizing interferometer. (b) High resolution filter composed of polarizing interferometers and birefringent elements.

VII. MATERIALS FOR BIREFRINGENT FILTERS

For the benefit of potential builders of birefringent filters, a brief discussion of available materials is given below. It must be emphasized that the list given is certainly far from complete. The author simply lists materials which have come to his attention and either have been successfully used, or look promising. Unfortunately, lack of time has prevented a really thorough search for suitable and available materials, and it would be surprising if some very useful ones had not been overlooked.

Some of the desirable properties of crystals for birefringent filters are a large value of μ with a small temperature coefficient; a high degree of hardness; chemical stability and insolubility in water; high transparency in the region of the spectrum for which the filter is to be used; and availability in large pieces of high optical quality.

For filters with band widths of 3 angstroms or more, quartz is an ideal material. It is excellent on all counts except for its rather small value of $\mu(=0.009)$. The birefringent elements of all the astronomical filters now in operation are made of quartz except for the final element of Lyot's filter, which is calcite.

Calcite would be excellent for elements of large n values if it were readily available in large sizes. Unfortunately it is so difficult to obtain that its general use in filters is probably impossible. While it is not as easily ground and polished as quartz, it presents no real difficulty. $\mu=0.17$.

Gypsum occurs naturally in large crystals and should be readily available. Its birefringence is similar to that of quartz, and it should be useful in the same places. Unfortunately, it is quite soft and might be difficult to polish. $\mu=0.009$.

Ammonium di-hydrogen phosphate has excellent optical characteristics, although it is sensitive to pressure and must be mounted with care. It is available in large sizes. Its optical working has proved rather difficult, though not impossible, and its high solubility in water necessitates careful protection from atmospheric moisture. $\mu=0.045$.

Ethylene diamine tartrate has promising optical characteristics accompanied by the disadvantages of high solubility in water and softness. The author knows of no attempts to polish it, but it would probably be quite difficult. It is available in large sizes. $\mu=0.084$.

Sodium nitrate has a larger μ-value than calcite, and should be useful for elements of large n-values. However, it is very soluble in water and difficult to work. At present it is not available in large sizes with the necessary homogeneity. $\mu=0.25$.

VIII. POLARIZING INTERFEROMETER FILTERS

An account of birefringent filters should not be closed without some mention of the polarizing interferometer, a device which has the effect of an impossibly thick birefringent element. It offers the possibility of filters of very high resolution with band widths in the range of hundredths or thousandths of an angstrom. The advantages of the polarizing over the usual forms of interferometers is in the possibility of an accurate and stable control of the wave-lengths of transmission maxima (by means of phase shifters) and a high light efficiency.

The essential feature of the polarizing interferometer is that the emerging light consists of two coherent sets of waves which differ in phase (because of path difference) and are polarized at right angles to each other. The effect is similar to that of a birefringent element, and a series of polarizing interferometers can be used exactly like a series of birefringent elements to construct a filter. The wave-length of the transmission band can be controlled with adjustable phase shifters, and interferometers can be sandwiched between birefringent elements to form split-element units.

The advantage of the polarizing interferometer over a simple birefringent element is that very large values of n can be obtained in a comparatively compact element. The saving in bulk may not be important, but the difficulty of obtaining birefringent material in very great thicknesses is significant. An element of calcite, for instance, must be about eleven times as thick as a path difference in glass. The principal disadvantage is the expense of construction common to all interferometers of the split amplitude class. The field is small for large values of n, and while it is theoretically quite simple to make a birefringent field compensator, it is impractical because the thickness of birefringent material required nullifies the advantage of compactness.

Many forms of polarizing interferometers are possible. One type which is well adapted for the construction of filters is shown at a, Fig. 7. It is a modified solid Michelson interferometer with a polarizing beam splitter. It consists of two glass prisms, A and B, with a very thin slip, b, of sodium nitrate (or other highly birefringent material) cemented between them with its optic axis normal to the surface. If the angles are properly chosen, the b-layer totally reflects the light vibrating in the plane of the drawing and transmits the light vibrating at right angles to it. A spacer element, C, introduces a path difference. Surfaces S and T are silvered or aluminized. Light which enters in the direction OS, emerges in the reverse direction, SO, in two components polarized at right angles, with a phase difference given by

$$2\pi n = 4\pi(\mu'/\lambda)dc\cos\varphi \qquad \text{(VIII.1)}$$

where μ' is the refractive index and φ is the angle of incidence on S and T. The prism P (constructed like A, B) has the double function of polarizing entering light and separating out the desired part of the emerging light. It is shown in an incorrect orientation for simplicity in drawing. Actually prism P is rotated about the OS direction, to bring its axis to an angle of 45° to that of prism AB. The transmission of the whole assembly for light emerging in the R direction is then

$$\tau = \sin^2\pi n. \qquad \text{(VIII.2)}$$

The remainder of the light emerges along SO.

The most serious difficulty in the construction of such an interferometer is the optical working and cementing of the b-layer to the required accuracy. The orientation of the S and T surfaces with respect to each other is not so critical, since a slight misalignment can be compensated by a thin wedge of birefringent material between prism P and the interferometer.

One method of using polarizing interferometers combined with birefringent elements in a filter is shown schematically at b, Fig. 7. Between each polarizer, P, and the following interferometer, I, is a b-element, which constitutes the m (for entering light) and q (for emerging light) components of a split element. The interferometer then takes the place of the p component. Between successive polarizers are purely birefringent split element units. The assembly includes 4 interferometers, 4 polarizing prisms, and 10 b-elements. The interferometers and b-elements should be equipped with phase shifters (not shown). As an example, the interferometers might have retardations of 245,760; 122,880; 61,440; 30,720; and the b-elements, retardations from 15360.5 to 30.5 at $\lambda = 5000$ angstroms. The system would transmit bands of about 0.01 angstrom effective width, spaced about 150 angstroms apart. Adjustment of the phase shifters will cause a selected band to scan the spectrum.

If the light transmitted by the filter is received on a photoelectric cell, its output gives a high resolution spectrophotometric curve of the entering light. Such a filter would be preferable to a grating spectrograph for spectrophotometric purposes, because, in spite of its small field (maximum usable φ about 0.0012 radian), it can be designed to transmit something like 1000 times as much light—a matter of considerable importance when such sharp bands are used, even in solar studies.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 39, NUMBER 3 MARCH, 1949

Diagnosing Types of Color Deficiency by Means of Pseudo-Isochromatic Tests

ALPHONSE CHAPANIS*
*Department of Psychology, The Johns Hopkins University, Baltimore, Maryland***
(Received September 27, 1948)

One hundred and twenty men (86 color-normal and 34 color-deficient) were tested with 5 pseudo-isochromatic tests of color vision: the American Optical Company, Ishihara, Meyrowitz, Boström, and Boström-Kugelberg tests. Each of the men was also examined with a Bausch and Lomb visual spectrophotometer to determine his limit of visibility in the long wave-length (red) end of the spectrum. Twelve of the color deficient subjects had reduced sensitivity to long wave-length radiation, 19 had normal sensitivity to long wave-length radiation, and 3 subjects could not be assigned to either group with certainty.

Diagnostic plates in three of the tests, the Boström, Meyrowitz, and Ishihara, were evaluated in terms of their ability to differentiate the two kinds of color deficient individuals. The Boström plates were found to be worthless for this purpose, the Meyrowitz plates somewhat better, and the Ishihara plates best. None of the plates in the other two tests was found to be diagnostic.

I. INTRODUCTION

In a recent article,[1] the author reported in detail on the comparative efficiency of five pseudo-isochromatic tests for separating color normal from color deficient individuals irrespective of the type of color deficiency involved. For some practical purposes, however, it is important to know the kind of color defect a person has, and a test which could clearly differentiate between the various types of color deficiency would be useful. Three of the five tests studied contain charts for differentiating between protanopes and deuteranopes. It is the purpose of this article to report on the diagnostic efficiency of these charts. They appear in the following tests:

1. C. G. Boström: *Tabulae Pseudo-Isochromaticae: Tavlor för Utrönande av Rubbningar i Färgsinnet.* Printed by Kifa, Stockholm, 1935. Distributed by the Nordiska Bokhandeln, Stockholm.

2. *Pseudo-Isochromatic Plates for Testing Color Perception.* (American Edition). Published by the E. B. Meyrowitz Surgical Instruments Company, Inc., New York, 1940.

3. S. Ishihara: *Tests for Colour-Blindness* (9th Edition). Printed by Kanehara and Company, Tokyo, 1940. Distributed by the C. H. Stoelting Company, Chicago, Illinois.

The two other tests used in this study were the following:

4. *Pseudo-Isochromatic Plates for Testing Color Perception.* Engraved and printed by the Beck Engraving Company, Inc., New York, 1940. Distributed by the American Optical Company, Southbridge, Massachusetts.

* The writer is indebted to Mrs. Mary Lamb, engineering aide, for her invaluable assistance in this study.

** The data of this study were collected while the author was Captain, Air Corps, in the Aero Medical Laboratory, Air Materiel Command, Wright Field, Dayton, Ohio.

[1] A. Chapanis, "A comparative study of five tests of color vision," J. Opt. Soc. Am. **38**, 626–649 (1948).

5. C. G. Boström and I. Kugelberg: *Tabulae Pseudo-Isochromaticae: Tavlor för Färgsinnesprövning.* Printed by Kifa, Stockholm, 1944. Distributed by the Nordiska Bokhandeln, Stockholm.

Although these two tests do not contain plates for differentiating between protanopes and deuteranopes, they will be referred to briefly in the later section of this paper.

II. SUBJECTS, APPARATUS, AND GENERAL PROCEDURES

Subjects

The subjects, apparatus, and procedures used in this experiment have been described in detail in the earlier paper.[1] Briefly, however, the subjects were 120 men selected from civilian and military personnel at Wright Field. Of these, 86 were color-normal and 34 color-deficient according to criteria defined in the earlier paper. This proportion of color-deficient to color-normal men is considerably greater than is characteristic of the population at large. There is reason to believe, however, that the color-deficient individuals in this study form a random sample of color-deficient men in general. The study was restricted to subjects whose previous medical history suggested a congenital type of color defect.

Apparatus and General Procedures

Each of the subjects was examined individually in a dark room according to the following schedule: (a) short medical history and interview; (b) visual acuity measurements; (c) test with the Bausch and Lomb visual spectrophotometer to determine the limit of visibility in the long wave-length (red) end of the spectrum; (d) examination with the five pseudo-isochromatic tests; and (e) examination with the Royal Canadian Navy Colour Vision Test Lantern.

The chart tests were placed on a table with a black felt background and were viewed at a distance of about 2½ feet. Illumination on the charts and table came from a MacBeth zenith sunlight (Type BX826 skylight) directly above the table. This light furnished 30 footcandles of illumination (of about 5500°K) incident on the test. The tests were tilted slightly to eliminate reflections from the light source. During this part of the examination, all other sources of illumination in the room were extinguished.

The order of presentation of the tests was randomized. There are 120 different permutations possible with these five tests and each subject was tested with a different order. No attempt was made to randomize the order of presentation of individual plates within the tests. The subjects were allowed unlimited time and all responses to each of the charts were noted. In no case, however, did the examiner supply any cue as to whether the response was correct. When a subject gave two or more responses to a chart, he was forced to state which one of the responses seemed to be clearest or "best" to him. These "best choice" responses are the ones which are scored and tabulated in the following data.

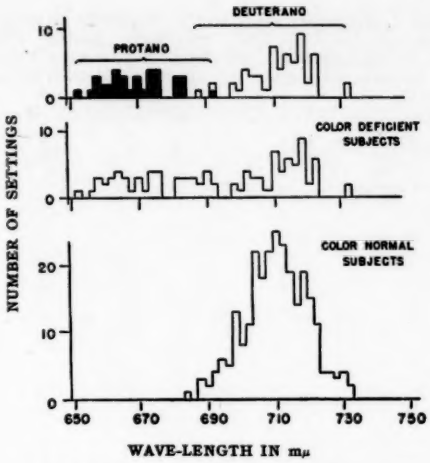

FIG. 1. Histograms of all wave-length settings (3 for each subject) made by 86 color-normal and 34 color-deficient subjects in locating the red end of the spectrum seen in a Bausch and Lomb visual spectrophotometer. The uppermost histogram does not include 3 individuals whose color defect could not be classified with certainty.

III. CLASSIFYING THE COLOR DEFICIENT SUBJECTS IN THIS EXPERIMENT

A Classification of Color Vision Defects

In order to have a common starting point, some classification of color vision defects should be agreed upon. Probably no single classification will satisfy all proponents of the various theories of color vision. This lack of agreement appears to be the result of insufficient data since most experiments on color deficiency have been limited to studies of a few individuals. There are two reasons for this: (a) color deficiency has a fairly low incidence in a normal population, and (b) the complete examination and description of a color-deficient individual with spectral apparatus is an extremely time-consuming procedure. Nonetheless, the accumulated experience of a hundred years has demonstrated that there are certain characteristics which seem to describe the color vision capacities of most people. This experience allows us to make a classification of color vision which most experts will probably agree with. It is a descriptive classification, independent of any color theory. A clear statement of this classification has been given by

TABLE I. Responses made to the diagnostic plates in the Boström test by 12 subjects who were classified protano (*A–L*), 19 subjects who were classified deuterano (*M–AE*), and 3 subjects who could not be classified with certainty (*AF–AH*).

Plate number	3	4	5	6	7	8	
Normal response	2	5	3	6	62	95	
Protano response	2	5	–	–	62	95	
Deuterano response	–	–	3	6	—	—	*D*-scores
Subject *A*	–	5	–	–	—	—	5
B	–	–	–	8	—	—	6
C	–	–	–	–	62	95	2
D	–	5	–	–	62	—	3
E	–	–	–	–	—	—	6
F	2	5	3	–	—	—	5
G	2	–	8	6	62	95	2
H	–	–	–	–	—	—	6
I	2	5	–	–	32	—	3
J	–	–	–	–	66	—	5
K	2	–	–	–	—	—	5
L	–	–	–	–	—	—	6
M	2	–	8	–	—	—	5
N	–	5	8	6	—	—	6
O	–	–	–	–	—	—	6
P	8	–	8	6	62	—	5
Q	9	5	3	6	6–	—	6
R	–	–	3	–	—	—	7
S	–	–	–	–	—	—	6
T	–	–	–	–	6–	—	5
U	–	–	–	–	—	—	6
V	–	–	8	6	—	99	6
W	–	–	–	6	—	—	7
X	–	–	–	–	—	—	6
Y	2	5	8	6	62	95	1
Z	–	–	–	–	—	—	6
AA	–	–	8	6	—	—	7
AB	–	–	–	6	—	—	7
AC	–	–	8	–	62	3–	4
AD	–	–	–	–	—	—	6
AE	8	5	8	–	—	58	5
AF	–	–	–	–	—	—	6
AG	–	–	–	–	—	—	6
AH	2	3	–	–	—	—	5

Hecht and Shlaer[2] and Judd[3] and will form the basis for the following discussion. In essence it states:

A. *Monochromats* are persons who confuse any part of the spectrum with any other part, and who can match any part with white.

B. *Dichromats* are persons who confuse large sections of the spectrum, who can match any part of the spectrum with a mixture of two primaries, and who can match certain hues (either blue-green, yellow, or extra-spectral purple) with white. Of these, the (a) *protanopes* confuse green, yellow, and red; they match blue-green and a purplish red in the extra-spectral region with white; and their spectral luminosity functions, with full daylight adaptation, are depressed in the red end of the spectrum. They are, therefore, frequently called "red-blind." The (b) *deuteranopes* also confuse green, yellow, and red; they match blue-green and a red-purple in the extra-spectral region with white; but they have spectral luminosity functions much like the normal. These are often, but inaccurately, called "green-blind." To be distinguished from both of these green-yellow-red confusers are the (c) *tritanopes* who confuse blue and green and match yellow with white. These are often called "violet-blind," or "blue-blind."

C. *Anomalous trichromats* are persons who confuse large sections of the spectrum, but who require mixtures of at least three primaries to match all parts of the spectrum. Their trichromatic mixtures, however, require unusual proportions of one of the primaries (as compared with mixtures made by the color normal) in order to achieve satisfactory matches. It is generally believed that anomalous trichromats occur as intermediates of various grades between dichromats and color normals. They are called (a) *protanomalous*, (b) *deuteranomalous*, or (c) *tritanomalous*, depending on which kind of dichromat they resemble.

D. *Trichromats* possess normal color vision. They require mixtures of three primaries in certain proportions to match the spectrum. Their brightness, hue, and saturation sensitivities to the spectral colors are known with some precision.

Types of Color Deficient Individuals in this Study.—Tritanopes, tritanomalous, and monochromats are extremely rare. There is no evidence from the data collected that any of the color deficient subjects in this study fell into any of these classifications. The rest of this paper, therefore, will be concerned with protanopes, deuteranopes, protanomalous, and deuteranomalous individuals.

Procedure for Diagnosing the Types of Color Deficiency

All of the color-deficient men in this experiment were red-green-yellow confusers. Some of them, however, were undoubtedly true dichromats and others were anomalous trichromats. It would have been desirable, for purposes of this study, to separate these two groups, but apparatus for making this kind of critical diagnosis was not available. It was possible, however, to separate the color-deficient into two groups, one containing subjects who saw the red end of the spectrum as shortened, and the other containing those who saw it normally. The former group included both protanopic and protanomalous individuals and, for lack of a better

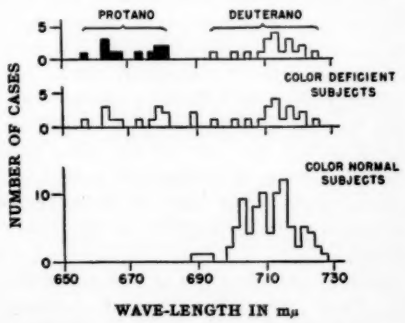

FIG. 2. Histograms of mean wave-length settings made by 86 color-normal and 34 color-deficient subjects in locating the red end of the spectrum seen in a Bausch and Lomb visual spectrophotometer. The uppermost histogram does not include 3 individuals whose color defect could not be classified with certainty.

[2] S. Hecht and S. Shlaer, "The color vision of dichromats: I. Wavelength discrimination, brightness distribution, and color mixture," J. Gen. Physiol. **20**, 57–82 (1936).

[3] D. B. Judd, "Facts of color-blindness," J. Opt. Soc. Am. **33**, 294–307 (1943).

TABLE II. Responses made to the diagnostic plates in the Meyrowitz test by 12 subjects who were classified protano (*A–L*), 19 subjects who were classified deuterano (*M–AE*), and 3 subjects who could not be classified with certainty (*AF–AH*).

Plate number	13	14	15	16	17	18	
Normal response	8	65	5	6	34	*RZ*	
Protano response	–	65	–	6	–4	*R–*	
Deuterano response	8	—	5	–	3–	*–Z*	*D*-scores
Subject *A*	8	—	–	6	8–	*B–*	5
B	8	—	5	6	84	*SZ*	6
C	8	—	–	6	—	—	5
D	8	—	5	6	24	87	5
E	8	—	–	6	—	—	5
F	8	—	–	6	34	*B2*	5
G	8	—	5	6	34	92	6
H	8	—	3	6	—	*–Z*	6
I	8	—	–	6	34	92	5
J	8	—	–	6	–4	57	4
K	8	—	–	6	84	*RZ*	4
L	8	—	–	6	34	82	5
M	8	—	5	6	3–	*RZ*	7
N	8	—	5	6	34	*–Z*	7
O	8	—	6	6	3–	*–Z*	7
P	8	—	5	6	8–	*–Z*	7
Q	8	—	5	6	34	*RZ*	6
R	8	—	5	6	3–	*–Z*	8
S	8	—	5	6	3–	*2Z*	8
T	8	—	5	–	84	*2Z*	7
U	8	—	5	6	3–	*–Z*	8
V	8	—	5	6	34	*BZ*	7
W	8	—	5	6	36	32	7
X	8	—	5	–	3–	*–Z*	9
Y	8	—	5	6	34	*BZ*	7
Z	8	—	5	6	3–	97	7
AA	8	—	5	–	3–	*–Z*	9
AB	8	—	5	–	84	*–Z*	7
AC	8	—	5	6	33	*–Z*	8
AD	8	—	5	–	34	*RZ*	7
AE	8	—	5	6	34	*RZ*	6
AF	8	—	–	6	—	–2	5
AG	8	—	5	6	34	*–Z*	7
AH	8	—	5	6	84	82	5

TABLE III. Responses made to the diagnostic plates in the Ishihara test by 12 subjects who were classified protano (*A–L*), 19 subjects who were classified deuterano (*M–AE*), and 3 subjects who could not be classified with certainty (*AF–AH*).

Plate number	22	23	24	25	
Normal response	26	42	35	96	
Protano response	–6	–2	–5	–6	
Deuterano response	2–	4–	3–	9–	*D*-scores
Subject *A*	–8	–2	–5	–6	1
B	26	42	35	36	3
C	–8	–2	–5	–6	1
D	–8	–2	–5	–6	1
E	—	—	—	—	4
F	26	42	45	–6	2
G	26	42	85	96	3
H	–6	–2	–5	–6	0
I	–6	–2	–5	–5	1
J	–8	–9	–3	–5	4
K	66	42	–5	–6	1
L	–8	–2	–5	–6	1
M	2–	4–	9–	9–	7
N	2–	4–	3–	9–	8
O	2–	4–	8–	9–	7
P	2–	4–	8–	8–	6
Q	2–	4–	3–	9–	8
R	2–	4–	5–	9–	7
S	2–	4–	8–	9–	7
T	2–	4–	5–	8–	6
U	2–	4–	5–	9–	7
V	9–	4–	33	9–	7
W	2–	4–	5–	3–	6
X	2–	4–	3–	8–	7
Y	26	42	39	96	5
Z	2–	4–	9–	3–	6
AA	8–	4–	8–	9–	6
AB	2–	4–	3–	—	7
AC	2–	4–	3–	9–	8
AD	2–	4–	3–	9–	8
AE	2–	4–	5–	9–	7
AF	2–	4–	—	—	6
AG	2–	4–	9–	3–	6
AH	–8	–2	–5	–6	1

inclusive term, will be referred to as the protano group. The latter group included both deuteranopic and deuteranomalous individuals and will be referred to as the deuterano group.

The Spectrophotometer Test.—The apparatus used for this purpose was a Bausch and Lomb visual spectromphotometer. The exit slit width was set at 2 millimeters and all tests were conducted with moderate light adaptation. The room illumination was of the order of 5 footcandles at the spectrophotometer. The test procedure was as follows: The wave-length control knob of the spectrophotometer was set in the region of 600 mμ. Subjects were instructed to turn the knob until they had located the point at which all light just faded from the exit slit and the field turned black. Forward and backward manipulation of the knob was permitted for maximum precision. Three such measurements were obtained for each subject.

Results of the Spectrophotometer Test.—The results of the spectrophotometer tests are shown in Figs. 1 and 2. Figure 1 contains histograms of *all* wave-length settings (three for each subject) made by the 86 color-normal and 34 color-deficient subjects. Figure 2 contains histograms of the *averages* of the three wave-length settings made by each subject. In general, there is a fair degree of consistency in the data. Slightly more than half of the subjects were able to make three settings which agreed within 10 mμ. Considering the conditions under which the test was administered—the novelty of the task and the inexperience of the subjects—the results are satisfactory. There was no difference between the various groups in the consistency with which they could locate the end of the spectrum.

For the purposes of this study, it seemed best to obtain two clear-cut groups of color-deficient subjects. The uppermost histograms in Figs. 1 and 2, accordingly, do not contain the data for three individuals whose color defect could not be classified with certainty. Two of these had average wave-length settings which fell at the lower end of the normal range (cf. Fig. 2), and, since it was not clear whether they should be classified as protano

or deuterano, their data are omitted. In addition, one other subject, whose settings were extremely variable, is omitted. The final result is two groups with virtually no overlapping between them. There is, further, almost no overlap between the settings made by the protanos and the color-normal; and, finally, the wave-length settings made by the deuteranos and the color-normal overlap completely. The 12 individuals shown by the shaded portion of these histograms form the protano group, and the 19 shown by the unshaded portion form the deuterano group, for purposes of this article. It might be pointed out, incidentally, that the two men at the upper end of the protano group (Fig. 2) have been used in another, more extensive study[4] so that their correct classification is virtually certain.

Can the Color Deficient Be Classified on the Basis of Their Thresholds for Long-Wave Energy?—The data shown by the middle histograms in Figs. 1 and 2 are interesting in their own right. The classification of color vision defects given earlier states that the protanope (or protanomalous) can be distinguished from the deuteranope (or deuteranomalous) by the fact that the former has reduced luminosity in the long wave-length end of the spectrum. So far as the author is aware, no large-scale surveys of *randomly selected* color-deviant individuals have ever been made to determine whether these two groups can, in fact, be differentiated in this respect. The data reported here are, unfortunately, too few to form any definite conclusion about the nature of the distribution. The middle histogram in Fig. 1, however, suggests that there are no discrete groups. If anything, the distribution appears to be continuous, although perhaps bimodal. This would be consistent with Murray's[5] belief that protanopia and deuteranopia do not represent distinct types, but merely the extremes in a continuous series.

There are several things wrong with the data collected in this part of the experiment. The intensity of the spectrophotometer lamp was not carefully controlled, nor were the viewing conditions specifiable. The subjects were inexperienced. The exit spectrum in the spectrophotometer did not contain equal amounts of energy at different wave-lengths. Even so, the results obtained are fairly consistent and highly suggestive. A further, more carefully controlled and larger study of this sort would probably yield some highly important data on this aspect of color deficiency.

According to the classification of color vision defects given earlier, one might expect that the protanomalous should be able to see long-wave radiation better than protanopes. Such a prediction rests on the assumption that the protanomalous have visual characteristics intermediate between protanopes and the color normal. Assuming that scores on the tests used in this experiment furnish some indication of the amount of color deficiency an individual has, this prediction is not confirmed. For the protano subjects, there is no correlation between the location of the end of the visible spectrum and scores on any of the color vision tests in this study taken either singly or all together. This is consistent with Wright's finding[6] that luminosity curves for the protanomalous appear to be coincident with luminosity curves for protanopes, and not intermediate between those for protanopes and color normal individuals. It is reminiscent of Murray's hypothesis[7] that shortening of the spectrum is an additional deficiency superimposed on, but unrelated to, a color vision defect.

IV. THE DIAGNOSTIC EFFICIENCY OF THE PSEUDO-ISOCHROMATIC TESTS

The Boström Test

Plates 3 through 8 of the Boström test are supposed to differentiate between protano and deuterano individuals. According to the test manual, if an examinee fails to read plates 5 and 6 correctly, he is protanopic or protanomalous; if he fails to read plates 3, 4, 7, and 8 correctly, he is deuteranopic or deuteranomalous to a serious extent. Detailed responses made to these six plates by the 34 subjects in this study are shown in Table I. A deuterano score (*D*-score) was computed for each subject by tallying the number of re-

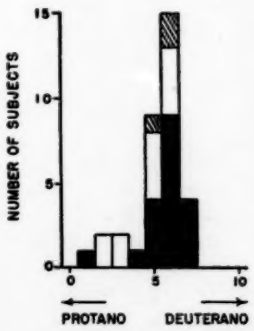

FIG. 3. Histogram of *D*-scores made by 34 color-deficient subjects on 6 diagnostic plates in the Boström test. The open areas represent 12 subjects who were classified protano; the solid areas, 19 subjects who were classified deuterano; and the striped areas, 3 subjects whose classification was uncertain on the basis of the spectrophotometer test. These data are taken from column 8 in Table I.

[4] A. Chapanis, "The dark adaptation of the color anomalous measured with lights of different hues," J. Gen. Physiol. **30**, 423–437 (1947).

[5] E. Murray, "Mass-testing of color vision: A simplified and accelerated technique," Am. J. Psych. **61**, 370–385 (1948).

[6] W. D. Wright, *Researches on Normal and Defective Color Vision* (The C. V. Mosby Company, St. Louis, 1947).

[7] E. Murray, "Color vision tests," pp. 275–282 in Glasser, O. (Ed.) *Medical Physics* (Year Book Publishers, Chicago, 1944).

sponses agreeing with those which should be read by the ideal deuterano subject (according to the test manual). High *D*-scores indicate conformity with the deuterano pattern of responses; low *D*-scores indicate non-conformity with the deuterano pattern, i.e., conformity with the protano pattern. Since there are 8 digits in these 6 plates, a protano score could have been computed by subtracting the *D*-score from 8. Since the protano and deuterano scores for any subject must add up to 8, however, the computation of a protano score would add nothing to the analysis.

D-scores for each subject are listed in the last column of Table I and are plotted in Fig. 3. It is immediately apparent from Fig. 3 that the diagnostic plates in the Boström test are not very efficient. They do not differentiate between color-deficient subjects who have shortened spectra and those who do not.

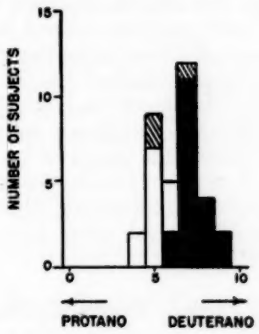

FIG. 4. Histogram of *D*-scores made by 34 color-deficient subjects on 6 diagnostic plates in the Meyrowitz test. The open areas represent 12 subjects who were classified protano; the solid areas, 19 subjects who were classified deuterano; and the striped areas, 3 subjects whose classification was uncertain on the basis of the spectrophotometer test. These data are taken from column 8 in Table II.

The Meyrowitz Test

According to the test manual, plates 13 through 18 of the Meyrowitz test are supposed to differentiate between protanopes and deuteranopes. The patterns of responses which should be made by both types of subjects are shown at the top of Table II, which lists the responses made by the 34 color-deficient subjects in this study. It is obvious from this table that plates 13 and 14 are worthless. Plate 13 is read by all of the color deficient individuals; plate 14 by none of them. Plate 16 is almost as bad; it is read by all of the protanos and almost all of the deuteranos. Plate 18 is only slightly better since most of the color deficient subjects read the plate correctly regardless of their classification. The other two Meyrowitz plates, 15 and 17, are considerably better. More than half of the protano and deuterano individuals gave responses to these plates according to the expected pattern.

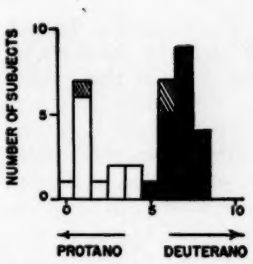

FIG. 5. Histogram of *D*-scores made by 34 color-deficient subjects on 4 diagnostic plates in the Ishihara test. The open areas represent 12 subjects who were classified protano; the solid areas, 19 subjects who were classified deuterano; and the striped areas, 3 subjects whose classification was uncertain on the basis of the spectrophotometer test. These data are taken from column 6 in Table III.

D-scores for these plates are shown in the last column of Table II and are plotted in Fig. 4. This chart shows that the Meyrowitz test is considerably better than the Boström for separating protanos from deuteranos. The deuterano subjects in this study earned *D*-scores of 6 or more; the protanos earned scores of 6 or less. Since there is overlapping of the *D*-scores for the two groups, however, it is questionable whether this test can be used reliably for diagnosing types of color deficiency.

The Ishihara Test

According to the test manual, Plates 22 through 24 of the Ishihara test can be used to differentiate between the "complete red-blind" and "complete green-blind." The performance of the 34 subjects on these plates is shown in Table III. As before, *D*-scores were computed, and Fig. 5 is a histogram of these scores. It is evident from this figure that the Ishihara test does the cleanest job of differentiating between color deficient individuals who have shortened spectra and those who do not. The protanos in this study earned *D*-scores of 4 or less; the deuteranos earned scores of 5 or more.

Digit-Misreading Errors.—In evaluating the efficiency of these diagnostic plates, it is evident that digit-misreading errors may markedly alter the results obtained. The *D*-scores in the last column of Table III were calculated with a literal scoring procedure. Thus, subject *A* responded "–8" to Ishihara plate 22. The "8" in this case is literally incorrect and so counts as a deuterano response. But it is also evident from the pattern of *A*'s responses to these plates that he is responding to those digits which only the protanope is supposed to see. Evidence was presented in the earlier paper[1] to suggest that digit-misreading errors are an indication of poor numeral design and are not necessarily related to color deficiency. The style of the Ishihara numerals was found to be better than some others, but normal individuals frequently

misread 3's as 8's; 3's as 9's; 9's as 8's; 6's as 8's; etc. These are precisely the sorts of errors which characterize the data in this table. These errors are made more frequently by the color deficient, to be sure, but they do raise a problem of scoring.

One way to circumvent this difficulty is to use a kind of clinical evaluation technique. If a subject responds only, or mostly, to those digits which the protanope is supposed to see, he may be classified as a protano no matter what he calls the digits. If a subject responds only, or mostly, to those digits which the deuteranope is supposed to see, he may be classified as a deuterano, again irrespective of what he reads. This simple scheme works surprisingly well. It fails to classify only 4 of the 34 subjects in this experiment, subjects *B*, *E*, *G*, and *Y*. Responses made by subjects *F*, *K*, and *AF* require a little study, but, in the main, they conform better to the protano than to the deuterano pattern.

As a general conclusion, if the Ishihara test classifies an examinee as being definitely protano or deuterano, the classification will almost certainly be correct. But not all color deficient subjects can be classified with such certainty. In short, if an experimenter is interested in selecting two clear-cut groups—a protano and a deuterano group—for some purpose, administering the Ishihara test is a simple way of doing it. Some color-deficient subjects will be "wasted" since they cannot be classified with certainty, but the two groups will almost certainly be correctly selected.

Related Studies.—The conclusions reached in the present study are supported by the findings of Hardy *et al.*[8] These investigators studied 22 color-deficient subjects (7 deuteranomalous, 5 protanomalous, 5 protanopes, and 5 deuteranopes) with the ninth edition of the Ishihara test. It may or may not be important to note incidentally that theirs was the British reprint, and not the original ninth edition. Of the 22 subjects, 17 (77 percent) were correctly classified, none were incorrectly classified, and the remaining 5 (23 percent) could not be classified with certainty. They also found, as in this study, that Ishihara plates 22–25 classify the deuteranomalous and deuteranopic somewhat more accurately than the protanomalous and protanopic.

A further estimate of the usefulness of these charts may be made from a study by Miles.[9] Among his group of 410 color-deficient subjects, 252 (61.5 percent) gave typical responses on Ishihara plates 22–25, i.e., they reported correctly all the protanopic digits and failed all the deuteranopic digits, or *vice versa*. An additional 45 (11.0 percent) saw correctly all the protanopic digits and some of the deuteranopic digits. Fifty-nine men (14.4 percent) had complete success with these charts, and the remaining 54 (13.2 percent) exhibited partial failure on both sets of digits. No independent means of verifying the type of color deficiency was used by Miles.

Does the Location of the Extra-Spectral Neutral Point Classify the Color Deficient?—The four diagnostic plates in the Ishihara test are a direct test of the question: Does this individual have a neutral point in the purplish red or in the red-purple region of the extra-spectral hues? But this implies that color-deficient subjects can, in fact, be differentiated in this respect. So far as the author is aware, there is no direct experimental evidence to support this assumption. It would be valuable to have accurate studies of a large number of randomly selected color deficient individuals to discover whether their extra-spectral neutral points form a single continuous distribution or two distributions, and, if the former, whether the distribution is unimodal or bimodal. Hecht and Shlaer[2] and Judd[10] have shown that the blue-green neutral point is worthless for diagnostic purposes because of the great amount of overlap between the neutral points of deuteranopes and protanopes. Similar data for the extra-spectral neutral point have apparently never been obtained.

Other Partially Diagnostic Plates

In addition to the specific plates designed to test for protanopia and deuteranopia, there was some evidence that certain other plates in the tests studied tended to favor either the protanos or the deuteranos. Among the *AO* plates, for example, plates 5 (Ishihara plate 2), 15, 16, 17, 18 (Meyrowitz plate 3), 21 and 22 tended to be harder for the protano subjects than the deuteranos. Similarly, plates 1, 2 (but not its equivalent Meyrowitz plate 8), 3, 4, 37, and 38 tended to be easier for the protanos and harder for the deuteranos. All of the protano subjects saw plates 2, 3, and 4 correctly, and all of the deuterano subjects saw plates 15 and 16 correctly.

Rowland[11] has recommended the use of *AO* plates 15 and 16 for differentiating deuteranopes from protanopes. In her study 52 deuteranopic subjects read these plates correctly, while only 6 of 21 protanopes (29 percent) were able to read them. She concludes that ". . . failure on charts #15

[8] LeG. H. Hardy, G. Rand, and M. C. Rittler, "The Ishihara test as a means of detecting and analyzing defective color vision," J. Gen. Psych. **36**, 79–106 (1947).

[9] W. R. Miles, "Color blindness in eleven thousand museum visitors," Yale J. Biol. Med. **16**, 59–76 (1943).

[10] D. B. Judd, "Standard response functions for protanopic and deuteranopic vision," J. Research Nat. Bur. of Stand. **33**, 407–437 (1944); also J. Opt. Soc. Am. **35**, 199–221 (1945).

[11] L. S. Rowland, "Selection of a battery of color vision tests," AAF School of Aviation Medicine, Randolph Field, Texas, Project No. 108, Report No. 1. (Feb. 19, 1943.)

and #16 may be considered a definite indication of a protanopic defect, but not all cases of such defect are detected." The results obtained here agree with her findings except that slightly more of the protanos (42 percent) were able to read these charts.

Among the Ishihara plates there was only one which seemed to be easier for the deuteranos as compared with the protanos. This was plate 2 (*AO* plate 5).

Among the Meyrowitz plates, plates 3 (*AO* plate 18) and 20 tended to favor the deuteranos, while plate 19 was a little easier for the protano subjects.

Among the Boström plates, 11 and 12 were considerably harder for the protano subjects than for the deuteranos. None of the protanos, in fact, saw these plates; whereas about 50 percent of the deuterano subjects saw them.

Among the Boström-Kugelberg plates, 10, 11, 12, 13, and 16 were considerably harder for the deuteranos than for the protanos. Because of the large number of plates on the Boström-Kugelberg test which seemed to favor them, protano subjects on the average received a higher score on this test than did the deuteranos. The average numbers of plates read correctly by the two groups were 2.3 and 1.0, respectively. This difference is significant at the 2 percent level of confidence when tested statistically. The overlap between the scores received by the two groups of subjects on this test is very great, however.

In general, none of these plates was sufficiently diagnostic to be useful for classifying color vision defects. It is also difficult to say how much importance should be attached to these slight differences, since Hardy *et al.*[8] have shown that several Ishihara charts may be differentially weighted for protanopic and deuteranopic vision by changing the quality of the illumination.

V. SUMMARY

Diagnostic plates in three pseudo-isochromatic tests, the Boström, Meyrowitz, and Ishihara, were evaluated in terms of their ability to differentiate color-deficient individuals who have reduced sensitivity to long-wave radiation from those who have normal sensitivity to long-wave radiation. The Boström plates were found to be worthless for this purpose, the Meyrowitz plates somewhat better, and the Ishihara plates best.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 39, NUMBER 3 MARCH, 1949

The Effect of Finite Aperture on the Fringe Intensity Distribution of Fabry Perot Interferometers*

FELIX E. GEIGER, JR.
University of Wisconsin, Madison, Wisconsin
(Received December 6, 1948)

The effect of finite aperture of a Fabry Perot interferometer on the fringe intensity distribution in the interferometer pattern is investigated. An equation is derived for a square etalon which illustrates clearly the decrease in fringe intensity for increasing angles of incidence. The computations show that this effect becomes significant only for high reflecting powers and comparatively large plate separations.

1. INTRODUCTION

THE intensity distribution in the interference fringes of a Fabry Perot interferometer, if the number of reflections is infinite, is given by the classical expression first derived by Airy,

$$I=1/(1+a^2\sin^2\delta), \qquad (1)$$

where δ is the phase difference between successively reflected beams, and $a^2=4r/d^2$, r being the reflecting power, and d the transmission coefficient of the silver film. Absorption is neglected.

However, if the number of reflections is finite, Eq. (1) becomes,

$$I_p=d^2[(1-r^p)^2+4r^p\sin^2p\delta/(1-r)^2+4r\sin^2\delta], \qquad (2)$$

where p is the number of interfering beams.

In practice,[1] it is assumed that p is very large and Eq. (1) holds.

Equation (1) represents an ideal state of affairs which does not exist in practice. Marked deviations in intensity from the theoretical intensities, as computed from Eq. (1), may take place in the fringe pattern as the angle of incidence (the angle which the incoming beam of light makes with the

* This work was done under contract N7onr285 T.O. No. 1.

[1] S. Tolansky, *High Resolution Spectroscopy* (Pitman Publishing Corporation, New York, 1947), Chapter VII, p. 98.

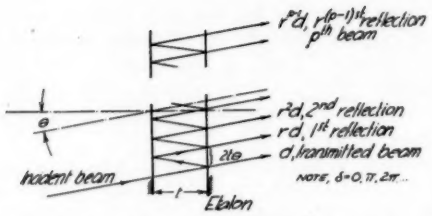

FIG. 1. Amplitudes of transmitted beam and multiply reflected beams. The first reflected beam is shifted by $2t\theta$. The phase difference 3δ, between successive beams is 0, π, 2π, etc.

normal to the etalon surface) is increased. This decrease in intensity of the interference fringes, with increasing angle of incidence, is wholly unaccounted for by Eq. (1) and becomes significant for reflecting powers greater than 80 percent and plate separations of 10 cm or more.

2. COMPUTATIONS

The effect of finite aperture on the fringe intensity distribution of the Fabry Perot will now be considered in detail.

If the angle of incidence is increased, the number of reflections decreases, and the effective aperture of the interferometer or, more precisely, of the optical system, diminishes for successive orders.[2] In order to simplify calculations, we assume that the etalon is a square of area A^2.

Summing amplitudes for a phase difference, δ, between consecutive beams, one gets for the resultant amplitude after $(p-1)$ reflections (Fig. 1)[3],

$$E_p = de^{i\delta}(1+re^{2i\delta}+\cdots+r^{p-1}e^{2(p-1)i\delta}). \quad (3)$$

Equation (2) is simply the square of Eq. (3). In order to take into account the finite, and, for successively reflected beams, ever decreasing aperture, one has to multiply each term of Eq. (3) by a correction factor.

Derivation of Correction Factor

Consider an etalon with plate separation t, and a beam of light incident upon it at an angle θ. Each beam is shifted by $2t\tan\vartheta$ after each reflection.[3] (Fig. 1.) Or for small angles, the shift will be $2t\vartheta$, and the number of reflections, $A/2t\vartheta$. The first transmitted beam makes full use of the available aperture of the etalon, the first reflected beam is shifted by the amount $2t\vartheta$ and, hence, suffers a decrease in intensity given by the ratio

[2] See reference 1, Chapter IX, p. 161.

[3] In Fig. 1 and elsewhere, whereas δ is the phase difference between strictly consecutive reflections, references to successively reflected beams, and to the nth and pth reflected beams apply, not to strictly consecutive reflections, but to consecutive beams reflected at the first plate and transmitted through the second to the camera.

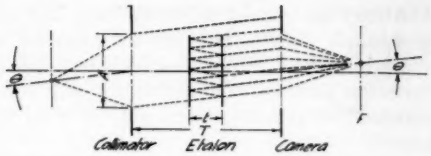

FIG. 2. Schematic diagram of a spectrograph crossed with the Fabry Perot etalon, showing the effect of the aperture of the camera on the number of beams brought to interference in the focal plane F.

of the areas,

$$[A^2 - A(2t\vartheta)]/A^2 = (1-2t\vartheta/A). \quad (4)$$

Consequently, the amplitude of the first reflected beam is decreased in the ratio

$$c_1 = (1-2t\vartheta/A)^{\frac{1}{2}}. \quad (5)$$

But p, the number of reflections,[3] is equal to $A/2t\vartheta$, and c_1 becomes very simply, $(1-1/p)^{\frac{1}{2}}$. Generally for the nth reflection we have

$$c_n = (1-n/p)^{\frac{1}{2}}, \quad (6)$$

or, if $n/p \ll 1$,

$$c_n = (1-n/2p), \quad n=1,\ 2,\ \cdots(p-1). \quad (7)$$

This approximation is quite poor for large n, but this will not affect our result materially, since for large n, the contributions of the reflected beams to the total intensity are very small.

Multiplying each term in Eq. (3) by the proper correction factor, c_n, we have,

$$E_p = de^{i\delta}(1+c_1re^{2i\delta}+\cdots+c_{p-1}r^{p-1}e^{2(p-1)i\delta}). \quad (8)$$

Summing the series, we obtain,

$$E_p = (de^{i\delta}/1-re^{2i\delta})\{(1-r^pe^{2pi\delta})-(re^{2pi\delta}/2p) \times[(1-r^{p-1})/(1-re^{2i\delta})-r^{p-1}(p-1)]\}. \quad (9)$$

For our purpose, it will be sufficient to consider the case when $\delta = 0, \pi, 2\pi, \cdots$etc., or the effect of finite aperture on the maximum fringe intensity. With $\delta = 0$, Eq. (9) becomes,

$$E_{max} = (d/1-r)\{(1-r^p)-(r/2p) \times[(1-r^{p-1})/(1-r)-(p-1)r^{p-1}]\}. \quad (10)$$

For $p \gg 1$, we may set $(p-1)$ equal to p,

$$E_{max} = (d/1-r)\{(1-r^p) \times[1-(r/2p(1-r))]+r^p/2\}. \quad (11)$$

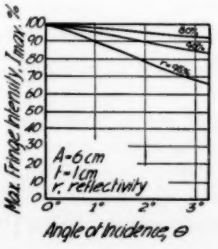

FIG. 3. Decrease of maximum fringe intensity for various reflectivities, and a plate separation of 1 cm, as the angle of incidence is increased.

Squaring, we get for the intensity,

$$I_{\max}=(d/1-r)^2\{(1-r^p)\times[1-(r/2p(1-r))]+r^p/2\}^2. \quad (12)$$

It is interesting to investigate the obvious limiting cases,

$$\lim_{p\to\infty} I_{\max}=1, \quad (13)$$

provided $d=(1-r)$, and

$$\lim_{r\to 1} I_{\max}=0. \quad (14)$$

Equation (13) may be verified at once from Eq. (12), in Eq. (14) one has to consider

$$\lim_{r\to 1}(1-r^p)/(1-r),$$

which, upon differentiation, becomes,

$$\lim_{r\to 1} pr^{p-1}=p,$$

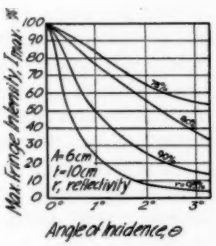

FIG. 4. Decrease of maximum fringe intensity for various reflectivities, and a plate separation of 10 cm, as the angle of incidence is increased.

and it is easily seen that Eq. (12) reduces to $-\frac{1}{2}+\frac{1}{2}=0$.

3. DISCUSSION

In an optical system, consisting of a Fabry Perot interferometer crossed with a spectrograph, the number of reflections will not be entirely determined by the aperture and the plate separation of the Fabry Perot alone. Figure 2 shows clearly that the number of reflections,[3] p, will be given by $(A-T\theta)/2t\theta$. It is assumed that the collimator lens, the camera lens, and the Fabry Perot have the same aperture. This diagram also shows that for this specific case the relative position of the etalon between the collimator and camera does not affect the number of reflections.

Several graphs have been drawn to illustrate Eq. (12). (See Figs. 3 and 4.)

Figure 5 shows an application of Eq. (12), and Figs. 3 and 4. For a certain setup, the interference pattern has been calculated to have maxima at approximately 0, 0.52, 0.73, 0.90, 1.04, 1.16 degrees. The Fabry Perot has a diameter of 6 cm, a reflecting power of 90 percent, and a plate separation of about 0.8 cm. Then the intensities of the maxima, neglecting absorption, will be 100, 98, 97, 96, 95, and 94 percent, respectively.

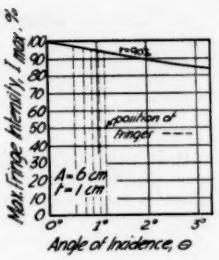

FIG. 5. In a given Fabry Perot interference pattern, maxima occur at approximately 0°, 0.52°, 0.73°, 0.90°, 1.04°, 1.16°. The reflecting power is 90 percent, the plate separation 1 cm, and $A=6$ cm. The graph shows how the intensity of the maxima decreases with decreasing order number.

4. ACKNOWLEDGMENT

The author wishes to express his appreciation for the interest of Dr. J. E. Mack, and an interesting discussion on this subject with Dr. K. W. Meissner.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 39, NUMBER 3 MARCH, 1949

The Color Perceptions of Deuteranopic and Protanopic Observers

DEANE B. JUDD
National Bureau of Standards, Washington, D. C.
(Received October 25, 1948)

It is well established that about 2 percent of otherwise normal human males are from birth confusers of red and green. There is considerable interest in the question: What do red-green confusers see? From a knowledge of the normal color perceptions corresponding to deuteranopic and protanopic red and green, we may not only understand better why color-blindness tests sometimes fail, and so be in a position to develop improved tests, but also the color-deficient observer may understand better the nature of his color-confusions and be aided to avoid their consequences. If an observer has normal trichromatic vision over a portion of his total retinal area and dichromatic vision over another portion, he may give valid testimony regarding the color perceptions characteristic of the particular form of dichromatic vision possessed by him. Preeminent among such observers are those born with one normal eye and one dichromatic eye. A review of the rather considerable literature on this subject shows that the color perceptions of both protanopic and deuteranopic observers are confined to two hues, yellow and blue, closely like those perceived under usual conditions in the spectrum at 575 mμ and 470 mμ, respectively, by normal observers. By combining this result with standard response functions recently derived for protanopic and deuteranopic vision, it has been possible to give quantitative estimates of the color perceptions typical of these observers for the whole range of colors in the Munsell Book of Color. These estimates take the form of protanopic and deuteranopic Munsell notations, and by using them it is possible not only to arrange the Munsell papers in ways which presumably appear orderly to red-green confusing dichromats but also to get immediately from the notations an accurate idea of the colors usually perceived in these arrangements by deuteranopes and protanopes much as the ordinary Munsell notations serve to describe the visual color perceptions of a normal observer.

I. INTRODUCTION

THE question, "What colors do color blind observers confuse?" is a very practical one, capable of objective solution by putting each color-blind observer to trial and noting his mistakes. This question has been given a fairly satisfactory general answer for protanopic and deuteranopic observers. The question, "What colors do color blind observers see?" is a more subtle one, involving a fine point in the theory of knowledge arising out of the fact that there is no common color language between the color blind and the normal eyed. Yet the essential facts have been ferreted out. It is one purpose of the present paper to review briefly the evidence by which these color perceptions have become qualitatively known. A second purpose is to indicate the basis of a comprehensive intertranslation, which is being published elsewhere,[1] between normal perceptions of surface colors on the one hand and protanopic and deuteranopic surface-color perceptions on the other. This intertranslation provides a method of expressing the color perceptions of the average red-green confuser in terms that are immediately comprehensible both to those trained in the interpretation of colorimetric coordinate systems and to the untrained alike. For most of the tabular and bibliographic material, reference must be had to the complete account.[1]

II. REVIEW OF THE LITERATURE

Persons born with one normal eye and one color-blind eye give us our most direct evidence of the color perceptions of the colorblind. Less reliable evidence may be obtained from persons who confuse red with green because of diseases of the eyes or optic nerves and who are therefore familiar with normal colors through past experience before becoming afflicted. And, finally, some information may be obtained from the peripheral parts of the normal retina which respond with confusions similar to those of deuteranopia. Any observer a portion of whose retinas yield normal color vision while another portion yields either the kind of red-green confusion characteristic of protanopia or that characteristic of deuteranopia can give valid testimony.

It is now a fairly well accepted view that red-green confusers see neutral colors normally and see chromatic colors of two hues approximately what normal observers call pure yellow and pure blue with little of an admixture of red and green. The literature has been reviewed to see whether this rough indication can be made more precise.

It is well known that the normal retina has a zone (perhaps 20 to 50°) from the fovea within which yellow, blue, black, and white are perceived much as at the fovea, but red-green distinctions are scarcely, if at all, possible. The luminosity function characterizing this zone is substantially normal, and therefore the properties of the zone approach those of deuteranopia. The precise hues of the yellowish and bluish colors seen by means of this zone are hard to determine. However, results by Hess, Baird, Dreher, and Goldman indicate that the yellow hue seen in the periphery corresponds to that of a spectrum stimulus of 567 to 575 mμ viewed foveally under customary observing conditions, and the blue hue corresponds to 460 to 471 mμ.

[1] D. B. Judd, "The color perceptions of deuteranopic and protanopic observers," J. Research Nat. Bur. of Stand. **41**, 247 (1948); RP1922.

Toxic agents and disease affecting the conducting elements of the visual mechanism (nerve elements and connections, optic nerve, and tract) cause a progressive lessening of the ability to distinguish red from green. The luminosity function is not essentially different from normal. Progressive red-green blindness therefore also corresponds to deuteranopia. According to Köllner,[2] a patient in the dichromatic state sees only yellow and blue in the spectrum, the yellow corresponding to the hue elicited in the normal eye by light of wave-length 575 mμ, the blue to the hue corresponding to that of wave-length 471 mμ.

A review of 37 cases of unilateral defects in color vision reported in the literature shows nine reports (Hippel, Holmgren, Hering, Hess, Nagel, Hayes, v. Kries, Goldschmidt, and Sloan) giving quantitative information regarding the hues perceived by protanopic and deuteranopic observers. The yellow hues reported correspond to spectrum stimuli ranging between 573 and 584 mμ; the blues range between 452 and 480 mμ. There is no significant difference between the protanopic and deuteranopic reports of hue. It is concluded that the wave-lengths 575 mμ and 470 mμ, respectively, represent these reports with no indication of significant individual variations. The most divergent report[3] of 584 and 452 mμ, respectively, was shown by auxiliary studies by means of Munsell papers[3] to be ascribable to an abnormally heavy ocular pigmentation.

In summary, it may be stated that evidence of three separate kinds combines to indicate that the hues perceived by protanopic and deuteranopic observers correspond closely to those perceived by normal-sighted observers to belong to spectrum stimuli of wave-length 575 mμ (yellow) and 470 mμ (blue).

III. AGREEMENT WITH THEORIES OF COLOR VISION

The evidence just reviewed that red-green confusers see yellow and blue of a certain kind does not prove that all red-green confusers have these perceptions of hue. The absence of conflicting evidence merely renders the validity of the conclusion highly probable. It is pertinent, therefore, to trace the influence exerted on color theories by these relatively few reports.

According to the original simple form of the three-components theory (red, green, violet), red-green blindness must come from a failure of either the red component (protanopia), or the green component (deuteranopia), from which it must follow that protanopes must see only mixtures of green and violet, and deuteranopes, only mixtures of red and violet. Printing-ink reproductions of the colors of the spectrum perceived by these types of red-green confusers, according to this theoretical view, have been published by Abney. Largely because of the failure of this simple theory to give a good explanation of the colors perceived in the peripheral portions of the normal retina or of those reported by observers having unilateral defects, the original form of the three-component theory is now generally regarded as insufficient. The modifications proposed

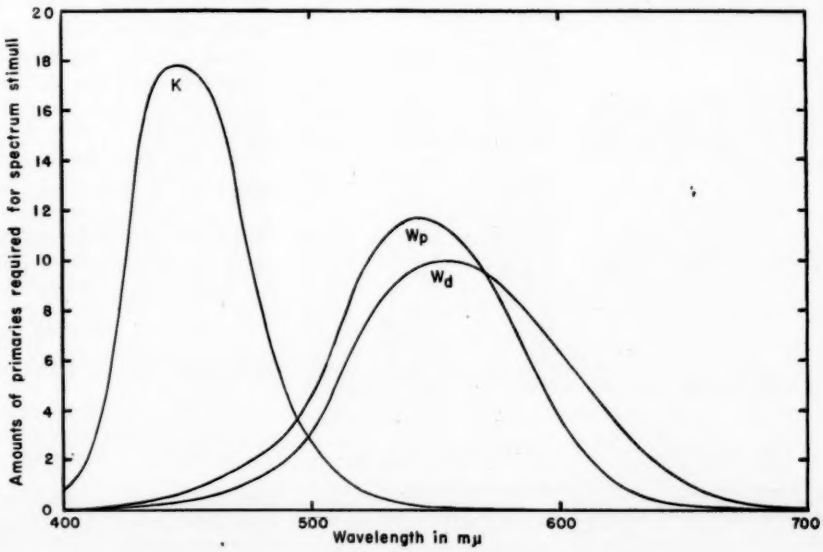

FIG. 1. Wave-length functions (K, W_p, W_d, see Eq. (1)) by means of which protanopic and deuteranopic red-green confusions may be computed. Any two colors having identical values of K and W_p are identical to the average protanope, and those having identical values of K and W_d are identical to the average deuteranope.

[2] H. Köllner, *Die Störungen des Farbensinnes* (Karger, Berlin, 1912).
[3] L. L. Sloan and L. Wollach, "Case of unilateral deuteranopia," J. Opt. Soc. Am. **38**, 502 (1948).

TABLE I. Deuteranopic and protanopic chromaticity coordinates for Munsell renotations of hues 5Y and 5PB.

Munsell renotation	Chromaticity coordinates w_d	w_p	Munsell renotation	Chromaticity coordinates w_d	w_p
5Y 9/12	0.8510	0.8433	*N* 1/ to 9/	0.4585	0.4652
5Y 9/10	0.7911	0.7827			
5Y 9/8	0.7296	0.7218			
5Y 9/6	0.6628	0.6569			
5Y 9/4	0.5955	0.5929			
5Y 9/2	0.5291	0.5305	5PB 9/2	0.4362	0.4457
5Y 8/12	0.8805	0.8731			
5Y 8/10	0.8181	0.8093			
5Y 8/8	0.7494	0.7408			
5Y 8/6	0.6764	0.6696	5PB 8/6	0.3615	0.3810
5Y 8/4	0.6025	0.5993	5PB 8/4	0.3973	0.4119
5Y 8/2	0.5325	0.5335	5PB 8/2	0.4325	0.4425
5Y 7/12	0.9125	0.9061			
5Y 7/10	0.8552	0.8466	5PB 7/10	0.2927	0.3214
5Y 7/8	0.7788	0.7693	5PB 7/8	0.3246	0.3489
5Y 7/6	0.7007	0.6927	5PB 7/6	0.3570	0.3768
5Y 7/4	0.6184	0.6139	5PB 7/4	0.3913	0.4065
5Y 7/2	0.5379	0.5383	5PB 7/2	0.4272	0.4377
5Y 6/12	0.9425	0.9377	5PB 6/12	0.2507	0.2856
5Y 6/10	0.8935	0.8859	5PB 6/10	0.2804	0.3107
5y 6/8	0.8231	0.8134	5PB 6/8	0.3096	0.3354
5y 6/6	0.7346	0.7251	5PB 6/6	0.3426	0.3638
5Y 6/4	0.6373	0.6314	5PB 6/4	0.3823	0.3984
5Y 6/2	0.5471	0.5468	5PB 6/2	0.4208	0.4319
5Y 5/12	0.9903	0.9894	5PB 5/12	0.2299	0.2675
5Y 5/10	0.9336	0.9279	5PB 5/10	0.2577	0.2909
5Y 5/8	0.8655	0.8562	5PB 5/8	0.2891	0.3176
5Y 5/6	0.7783	0.7677	5PB 5/6	0.3242	0.3476
5Y 5/4	0.6667	0.6588	5PB 5/4	0.3662	0.3838
5Y 5/2	0.5569	0.5556	5PB 5/2	0.4106	0.4226
			5PB 4/12	0.2017	0.2426
			5PB 4/10	0.2282	0.2653
5Y 4/8	0.9153	0.9081	5PB 4/8	0.2596	0.2919
5Y 4/6	0.8200	0.8092	5PB 4/6	0.2997	0.3261
5Y 4/4	0.7061	0.6962	5PB 4/4	0.3442	0.3642
5Y 4/2	0.5774	0.5741	5PB 4/2	0.3956	0.4091
			5PB 3/12	0.1606	0.2065
			5PB 3/10	0.1886	0.2305
			5PB 3/8	0.2223	0.2594
5Y 3/6	0.8839	0.8745	5PB 3/6	0.2605	0.2919
5Y 3/4	0.7373	0.7261	5PB 3/4	0.3104	0.3346
5Y 3/2	0.5899	0.5854	5PB 3/2	0.3729	0.3890
			5PB 2/12	0.1213	0.1709
			5PB 2/10	0.1459	0.1925
			5PB 2/8	0.1793	0.2215
			5PB 2/6	0.2247	0.2604
5Y 2/4	0.8380	0.8266	5PB 2/4	0.2833	0.3107
5Y 2/2	0.6149	0.6182	5PB 2/2	0.3564	0.3740
			5PB 1/10	0.0998	0.1506
			5PB 1/8	0.1314	0.1785
			5PB 1/6	0.1739	0.2157
			5PB 1/4	0.2337	0.2671
5Y 1/2	0.7392	0.7267	5PB 1/2	0.3127	0.3355

by Fick and by Ladd-Franklin both provide an explanation for protanopic and deuteranopic vision in which the hues of the colors perceived are yellow and blue, respectively.

According to the opponent-colors theory proposed by Hering, there can be only one form of red-green confusion, that resulting from the failure of the red-green process. This leaves the yellow-blue and the black-white processes intact. The form thus predicted is deuteranopia, and the hues are indicated correctly.

More recent theories of vision take account of the possibility of regroupings of elementary processes at more than one stage of the visual mechanism and are called stage theories. The most prominent of these theories (v. Kries, Schrödinger, Adams, Müller) have the property in common that the final stage is very like the Hering "opponent-colors" formulation.

In summary, all theories of vision with any pretense of being complete provide for the perception of yellow and blue by red-green confusers.

IV. DERIVATION OF DEUTERANOPIC AND PROTANOPIC MUNSELL NOTATIONS

Since there is no experimental or theoretical evidence against the view that yellow and blue are the hues perceived by red-green confusers, it is proposed to build a method of designating protanopic and deuteranopic perceptions of surface colors on this basis. It has been shown[4] that the color confusions of red-green-blind observers can be found from three numbers (K, W_p, W_d) related to the tri-stimulus values (X, Y, Z) of the ICI standard observer by the transformation equations

$$\left.\begin{aligned} K &= Z \\ W_p &= -0.460X + 1.359Y + 0.101Z \\ W_d &= Y \end{aligned}\right\}. \qquad (1)$$

Any two colors having identical values of K and W_p are identical to the average protanope; those having identical values of K and W_d are identical to the average deuteranope. Figure 1 shows in arbitrary units the values of K, W_p, and W_d for all parts of a spectrum of unit irradiance per unit wave-length.

For the ICI observer luminous directional reflectance is customarily expressed relative to that (Y_0) of magnesium oxide illuminated and viewed under geometrical conditions identical to those used for the specimen itself. By analogy we take for deuteranopic luminous directional reflectance

$$W_d/(W_d)_0 = Y/Y_0, \qquad (2)$$

and for protanopic luminous directional reflectance

$$\frac{W_p}{(W_p)_0} = \frac{-0.460X + 1.359Y + 0.101Z}{-0.460X_0 + 1.359Y_0 + 0.101Z_0}. \qquad (3)$$

But since for magnesium oxide illuminated by standard source C (representative of average day-

[4] D. B. Judd, "Standard response functions for protanopic and deuteranopic vision," J. Research Nat. Bur. of Stand. 33, 407 (1944); RP1618; also J. Opt. Soc. Am. 35, 199 (1945).

light) it is customary to take $X_0=0.9804$, $Y_0=1.0000$, and $Z_0=1.1512$, protanopic luminous directional reflectance for this illuminant may be written simply as $0.9733W_p$.

Similarly, by analogy to the customary practice of specifying the chromatic aspect of a color by chromaticity coordinates (x, y, z), we define deuteranopic chromaticity coordinates (w_d, k_d) as

$$\begin{aligned} w_d &\equiv W_d/(W_d+K)=Y/(Y+Z) \\ &\quad =y/(y+z)=y/(1-x), \\ k_d &\equiv K/(W_d+K)=Z/(Y+Z)=z/(y+z) \\ &\quad =(1-x-y)/(1-x). \end{aligned} \tag{4}$$

And we define protanopic chromaticity coordinates (w_p, k_p) as

$$\begin{aligned} w_p &\equiv \frac{W_p}{W_p+K}=\frac{-0.460X+1.359Y+0.101Z}{-0.460X+1.359Y+1.101Z} \\ &\quad =\frac{-0.561x+1.258y+0.101}{-1.561x+0.258y+1.101}, \\ k_p &\equiv \frac{K}{W_p+K}=\frac{Z}{-0.460X+1.359Y+1.101Z} \\ &\quad =\frac{1.000-x-y}{-1.561x+0.258y+1.101}. \end{aligned} \tag{5}$$

Since each pair of dichromatic coordinates sums to unity by definition, only one of each need be used. We accordingly select w_d and w_p as indications of the deuteranopic and protanopic "warm" quality, respectively. According to the data and views already summarized, it is legitimate to regard them more precisely as indications of "yellowness." Values of w_d and w_p together with the reflectance $W_d/(W_d)_0$, and $W_p/(W_p)_0$ have been computed from the known tristimulus values for all of the chips of the 1929 Munsell Book of Color[5] by means of Eqs. (2), (3), (4), and (5) (see complete paper[1]).

Evaluation of these quantities serves to indicate which chips of the Munsell Book of Color would be nearly of the same appearance when viewed by an average protanope or deuteranope, but they do not give an easily understood indication of the color perceptions themselves. The basic premises for which evidence has already been summarized are, however, that red-green confusers perceive black, gray, and white normally, that their yellow corresponds closely to what an observer having normal color vision sees in the spectrum at 575 mμ, and that their blue corresponds closely to 470 mμ. These facts may be conveniently expressed in the Munsell color system by giving for each set of values of W/W_0 and w the Munsell equivalent of the purple-blue (5PB) or yellow (5Y) hue; that is, we identify each color perceived by a red-green confuser by giving the Munsell notation of the color of Munsell hue 5Y or 5PB with which he would confuse it. These Munsell hues correspond closely to those of the spectrum at 470 and 575 mμ, respectively, and since the exact hues corresponding to these portions of the spectrum depend somewhat on the observing conditions, more precise fitting (as by 6PB and 7Y) is probably not warranted.

The deuteranopic and protanopic Munsell renotations based in this way resulted from the following steps: (1) From Table I of the OSA subcommittee report on the spacing of the Munsell colors[6] were read the chromaticity coordinates (x, y) for enough colors of the hues 5Y and 5PB to cover the range of the Munsell papers. (2) From

[5] *Munsell Book of Color* (Library and Pocket Editions) (Munsell Color Company, Baltimore, 1929).

TABLE II. Deuteranopic arrangement.

Munsell book notation	Deuteranopic reflectance $W_d=Y$	Deuteranopic chromaticity coordinate w_d	Munsell renotations Deuteranopic	Munsell renotations Normal
5Y 8/12	0.5706	0.8911	5Y 7.9/12.2	5.5Y 7.9/12.3
5Y 8/10	.5784	.8578	5Y 7.9/11.0	5.5Y 7.9/11.1
10Y 8/ 8	.6035	.7918	5Y 8.1/ 9.2	0.5GY 8.1/ 9.1
5Y 8/ 8	.5794	.7660	5Y 7.9/ 8.2	5.5Y 7.9/ 8.2
5Y 8/ 8	.5732	.7606	5Y 7.9/ 8.1	5.5Y 7.9/ 8.3
5GY 8/ 8	.6022	.7464	5Y 8.1/ 7.9	4.5GY 8.1/ 3.7
10YR 8/ 8	.6049	.7064	5Y 8.1/ 6.9	0.5Y 8.1/ 7.6
5Y 8/ 6	.5784	.6681	5Y 7.9/ 5.7	4.5Y 7.9/ 5.8
10Y 8/ 6	.5996	.6664	5Y 8.0/ 5.7	0.5GY 8.0/ 5.7
5GY 8/ 6	.5929	.6605	5Y 8.0/ 5.5	5.5GY 8.0/ 6.2
10YR 8/ 6	.6114	.6263	5Y 8.1/ 4.7	9.0YR 8.1/ 5.5
5Y 8/ 4	.5651	.5851	5Y 7.8/ 3.4	4.0Y 7.8/ 3.5
10Y 8/ 4	.5997	.5827	5Y 8.0/ 3.4	1.0GY 8.0/ 3.4
10GY 8/ 6	.5793	.5720	5Y 7.9/ 3.1	9.5GY 7.9/ 5.1
5GY 8/ 4	.6041	.5719	5Y 8.1/ 3.2	6.0GY 8.1/ 3.6
10YR 8/ 4	.5972	.5579	5Y 8.0/ 2.8	9.0YR 8.0/ 3.4
10GY 8/ 4	.6181	.5406	5Y 8.2/ 2.3	0.5G 8.2/ 4.0
5YR 8/ 4	.5970	.5311	5Y 8.0/ 2.0	4.0YR 8.0/ 3.5
10R 8/ 4	.6115	.5156	5Y 8.1/ 1.6	10.0R 8.1/ 3.7
5Y 8/ 2	.5570	.5141	5Y 7.8/ 1.5	3.5Y 7.8/ 1.6
5GY 8/ 2	.5851	.5107	5Y 8.0/ 1.5	6.0GY 8.0/ 1.6
5G 8/ 6	.5595	.4964	5Y 7.8/ 1.1	6.5G 7.8/ 4.7
5YR 8/ 2	.5917	.4949	5Y 8.0/ 1.0	4.0YR 8.0/ 2.0
5G 8/ 4	.5881	.4899	5Y 8.0/ 0.9	5.5G 8.0/ 3.3
5G 8/ 2	.5694	.4847	5Y 7.9/ 0.8	4.5G 7.9/ 2.1
5R 8/ 4	.5862	.4834	5Y 8.0/ 0.7	3.5R 8.0/ 3.6
10G 8/ 2	.5965	.4806	5Y 8.0/ 0.6	5.0G 8.0/ 2.0
N 8/	.5751	.4700	5Y 7.9/ 0.4	2.5GY 7.9/ 0.3
5R 8/ 2	.6021	.4699	5Y 8.1/ 0.4	2.0R 8.1/ 2.2
10RP 8/ 6	.6251	.4693	5Y 8.2/ 0.4	8.5RP 8.2/ 3.6
10RP 8/ 4	.6107	.4678	5Y 8.1/ 0.3	9.5RP 8.1/ 2.7
5BG 8/ 2	.6082	.4656	5Y 8.1/ 0.3	1.0BG 8.1/ 2.0
5RP 8/ 2	.6435	.4633	5Y 8.3/ 0.2	8.0RP 8.3/ 1.9
5RP 8/ 6	.6058	.4611	5Y 8.1/ 0.1	6.5RP 8.1/ 3.4
5RP 8/ 4	.6050	.4588	5Y 8.1/ 0.1	5.5RP 8.1/ 2.8
10BG 8/ 2	.6258	.4557	5PB 8.2/ 0.4	5.0BG 8.2/ 1.9
5B 8/ 2	.6179	.4361	5PB 8.2/ 1.7	4.0B 8.2/ 2.2
5P 8/ 2	.5897	.4330	5PB 8.0/ 1.8	5.0P 8.0/ 2.5
10P 8/ 4	.6182	.4289	5PB 8.2/ 2.2	7.5P 8.2/ 3.5
5B 8/ 4	.6025	.4272	5PB 8.1/ 2.3	4.0B 8.1/ 2.9
10B 8/ 2	.6514	.4170	5PB 8.3/ 3.1	1.0PB 8.3/ 3.0
5PB 8/ 2	.6380	.4139	5PB 8.3/ 3.2	5.0PB 8.3/ 3.2
10PB 8/ 2	.6499	.4099	5PB 8.3/ 3.4	7.5PB 8.3/ 3.4
P 8/ 4	.5940	.4004	5PB 8.0/ 3.9	4.5P 8.0/ 4.8

(For continuation of this table see reference 1.)

TABLE III. Protanopic arrangement.

Munsell book notation	Protanopic reflectance 0.9733 W_p	Protanopic chromaticity coordinate w_p	Munsell renotations Protanopic	Munsell renotations Normal
10BG 8/ 2	0.6464	0.4705	5Y 8.3/ 0.2	5.0BG 8.2/ 1.9
5B 8/ 2	.6391	.4511	5PB 8.3/ 1.3	4.0B 8.2/ 2.2
10B 8/ 2	.6732	.4317	5PB 8.4/ 2.8	1.0PB 8.3/ 3.0
5PB 8/ 2	.6543	.4267	5PB 8.3/ 3.2	5.0PB 8.3/ 3.2
10PB 8/ 2	.6608	.4205	5PB 8.4/ 3.5	7.5PB 8.3/ 3.4
10Y 8/ 8	.5773	.7889	5Y 7.9/ 9.3	0.5GY 8.1/ 9.1
5GY 8/ 8	.5991	.7505	5Y 8.0/ 8.3	4.5GY 8.1/ 8.7
5GY 8/ 6	.5914	.6659	5Y 8.0/ 5.9	5.5GY 8.0/ 6.2
10Y 8/ 6	.5817	.6657	5Y 8.0/ 5.8	0.5GY 8.0/ 5.7
5Y 8/ 6	.5451	.6609	5Y 7.7/ 5.7	4.5Y 7.9/ 5.8
10YR 8/ 6	.5632	.6134	5Y 7.8/ 4.3	9.0YR 8.1/ 5.5
10GY 8/ 6	.5987	.5866	5Y 8.0/ 3.6	9.5GY 7.9/ 5.1
10Y 8/ 4	.5873	.5842	5Y 8.0/ 3.5	1.0GY 8.0/ 3.4
5GY 8/ 4	.6044	.5786	5Y 8.1/ 3.4	6.0GY 8.1/ 3.6
10GY 8/ 4	.6362	.5545	5Y 8.2/ 2.7	0.5G 8.2/ 4.0
10YR 8/ 4	.5656	.5512	5Y 7.9/ 2.5	9.0YR 8.0/ 3.4
5YR 8/ 4	.5614	.5225	5Y 7.8/ 1.6	4.0YR 8.0/ 3.5
5G 8/ 6	.5954	.5187	5Y 8.0/ 1.5	6.5G 7.8/ 4.7
5GY 8/ 2	.5851	.5174	5Y 8.0/ 1.5	6.0GY 8.0/ 1.6
5G 8/ 4	.6141	.5075	5Y 8.1/ 1.3	5.5G 8.0/ 3.3
10R 8/ 4	.5725	.5059	5Y 7.9/ 1.2	10.0R 8.1/ 3.7
5G 8/ 2	.5854	.4984	5Y 8.0/ 0.9	4.5G 7.9/ 2.1
10G 8/ 2	.6133	.4943	5Y 8.1/ 0.8	5.0G 8.0/ 2.0
5YR 8/ 2	.5731	.4937	5Y 7.9/ 0.8	4.0YR 8.0/ 2.0
5BG 8/ 2	.6281	.4803	5Y 8.2/ 0.5	1.0BG 8.1/ 2.0
5R 8/ 4	.5565	.4772	5Y 7.8/ 0.4	3.5R 8.0/ 3.6
N 8/	.5741	.4764	5Y 7.9/ 0.3	2.5GY 7.9/ 0.3
5R 8/ 2	.5857	.4698	5Y 8.0/ 0.2	2.0R 8.1/ 2.2
5RP 8/ 2	.6317	.4654	N 8.2/	8.0RP 8.3/ 1.9
10RP 8/ 4	.5905	.4662	N 8.0/	9.5RP 8.1/ 2.7
10RP 8/ 6	.5958	.4641	5PB 8.0/ 0.1	8.5RP 8.2/ 3.6
5RP 8/ 6	.5839	.4586	5PB 8.0/ 0.6	6.5RP 8.1/ 3.4
5RP 8/ 4	.5874	.4582	5PB 8.0/ 0.7	5.5RP 8.1/ 2.8
5B 8/ 4	.6310	.4452	5PB 8.2/ 1.8	4.0B 8.1/ 2.9
5P 8/ 2	.5877	.4388	5PB 8.0/ 2.2	5.0P 8.0/ 2.5
10P 8/ 4	.6081	.4316	5PB 8.1/ 2.8	7.5P 8.2/ 3.5
5P 8/ 4	.5903	.4054	5PB 8.0/ 4.5	4.5P 8.0/ 4.8
5Y 8/12	.5208	.8848	5Y 7.6/11.8	5.5Y 7.9/12.3
5Y 8/10	.5292	.8501	5Y 7.6/10.8	5.5Y 7.9/11.1
5Y 8/ 8	.5393	.7579	5Y 7.7/ 8.3	5.5Y 7.9/ 8.3
5Y 8/ 8	.5339	.7525	5Y 7.7/ 8.2	5.5Y 7.9/ 8.2
10YR 8/ 8	.5451	.6902	5Y 7.7/ 6.4	0.5Y 8.1/ 7.6
5Y 8/ 4	.5426	.5818	5Y 7.7/ 3.4	4.0Y 7.8/ 3.5
5Y 8/ 2	.5455	.5156	5Y 7.7/ 1.5	3.5Y 7.8/ 1.6

(For continuation of this table see reference 1.)

these chromaticity coordinates the dichromatic coordinates w_d and w_p were computed, according to Eqs. (4) and (5). (3) Families of curves were then plotted from these data on each of two graphs, one with w_d as the abscissa, the other with w_p as the abscissa, the ordinate being Munsell chroma, each curve then showing the variation of Munsell chroma for hue 5Y or 5PB with dichromatic coordinate (w_d or w_p) for some one Munsell value. (4) From the reflectance relative to magnesium oxide, W/W_0, (found from Eqs. (2) and (3)) the Munsell value was read from Table II of the subcommittee report.[6] (5) Finally, by interpolation among the curves of the corresponding family according to this Munsell value, the Munsell hue (whether 5Y or 5PB) was determined, and the Munsell chroma corresponding to the dichromatic coordinate, w_d or w_p, was read with an uncertainty of about 0.1 chroma step.

Munsell dichromatic renotations (hue value/chroma) were found in this way for the chips (about 400) of the 1929 Munsell Book of Color. Table I shows the chromaticity coordinates w_d and w_p obtained in step (b); and Tables II and III show the Munsell dichromatic renotations for the chips of value 8/ compared to the normal renotations. Similar data for other Munsell values are given in the complete paper.[1] Comparison of the dichromatic Munsell renotations with the corresponding normal Munsell renotation indicates in a precise and detailed way the difference between the surface-color perceptions of the normal observer and those of the average deuteranope and average protanope, respectively.

It is suggested that persons who have been found to have either protanopic or deuteranopic vision by the usual tests might find it interesting and possibly instructive to lay out the Munsell papers in the order given in Tables II or III, as the case may be, and see to what degree the Munsell dichromatic notation accords with their own perceptions. To the extent that the accord is good, the person will discover how well his type of vision is known, and he may be encouraged to study the basis of the Munsell dichromatic notation and so gain an additional insight into the relation of his own vision to that possessed by the majority. If the Munsell dichromatic notation fails to accord with his perceptions in ways not to be corrected by viewing the papers through bluish or brownish goggles, it is likely that the observer will find that he has been mistyped by careless administration or interpretation of a routine test, and a retest would be in order. Should the observer find that he is indeed a red-green confusing dichromat, he would have a chance to contribute to the knowledge of color perceptions by studying and reporting the nature and degree of any contradictions.

[6] S. M. Newhall, D. Nickerson, and D. B. Judd, "Final report of the O.S.A. subcommittee on the spacing of the Munsell colors," J. Opt. Soc. Am. **33**, 385 (1943).

Spectrographic Analysis of High Cobalt (Vitallium) Alloys

Y. T. SIHVONEN, D. L. FRY, R. E. NUSBAUM,* AND R. R. BAUMGARTNER**
Research Laboratories Division, General Motors Corporation, Detroit, Michigan
(Received October 11, 1948)

A spectrochemical method of analysis used for control purposes in casting Vitallium is described in this paper. The immediate need for such a method was so urgent that only a small amount of development time was possible; control analyses were being made within three days of the time the problem was submitted. Improvements were made in the method in the course of the work. Reproducibility and accuracy data are given in the paper. This problem illustrates that a great deal of analytical work can be done with a spectrograph even though accurate comparison standards are not available.

SOME time ago our laboratory found it necessary to develop a rapid method of analyzing Vitallium for the six elements shown in Table I. This paper describes the spectrographic procedure adopted, and in so doing also shows how it is possible to set up analytical curves for control work even though accurate comparison standards are not available.

APPARATUS

The apparatus used in obtaining the data discussed in this paper is the regular spectrographic equipment of the laboratory and has been referred to in previous papers.[1,2]

STANDARD SAMPLES

As the casting and physical testing of high temperature alloys were already under way when the analysis problem was submitted, it was necessary to obtain a satisfactory analytical method as soon as possible so as not to impede the metallurgical project. To this end some one-inch rounds, formed previously for tensile tests and already analyzed chemically, were used for standard samples.

These standards were not entirely satisfactory as will be pointed out later. Yet, had the work been delayed until adequate standards were available, much of the value of making the analyses would have been lost. Although the early results were not accurate, variations from one sample to another could be reported. Later, after a large number of chemical checks had been made, the analytical curves were drawn through the statistical center of the concentration *versus* line-intensity ratio points.

PREPARATION OF SAMPLES

Because the standards were one-inch rounds, the unknowns also were of this size. This did not necessarily mean that the best results could be obtained on these samples. In fact, the metallurgists suggested that some segregation in such a casting was highly possible, and later work has shown that probably a $\frac{1}{4}''$ pin chill cast would have given better results.

It was found that for the one-inch rounds the best reproducibility was obtained with a highly polished surface. One end of each sample was ground on a machine-driven, 240 grit emery belt.

EXCITATION OF THE SPECTRUM

It has been found in the past that satisfactory operation in spark analyses can be recognized to some extent by observing the constancy and intensity of the spark and the way it burns the electrode surfaces. Desirable conditions are that the spark be stable and that the electrodes be fused only slightly and quite regularly. These requirements were most nearly fulfilled by using 0.021 microfarad capacity, 12 ohms, and no added inductance with $\frac{1}{16}''$ analytical gap in a Feusner spark circuit. With one exception, to be referred to later, small variations in these constants caused no determinable variation in the line ratios.

Spectrum lines of the same relative intensity, which did not interfere with lines of other elements which might be present in the sample, and which were close together, were sought for analysis lines. Because of the scarcity of appropriate lines, silicon

TABLE I.

Element	Range in percent	Spectrum lines	Col. 4	Col. 5	Col. 6
Molybdenum	5.35– 6.00	2644.353	5.62%	81	4.7%
Manganese	0.30– 0.40	2618.143	9.30	80	6.5
Iron	0.30– 0.50	2599.396	5.54	92	4.5
Chromium	27.0 –28.5	2571.741	4.32	81	4.0
Silicon	0.30– 0.40	2516.123	11.17	81	10.3
Nickel	1.60– 3.20	2394.501	3.73	92	3.5
Cobalt	63.0 –66.0	2497.501			

Column 4—Standard deviation on repeat exposures.
Column 5—Number of repeat exposures.
Column 6—The average deviation between values obtained by wet chemical methods and those obtained by spectrographic methods.

* Now at North American Aviation, Inglewood, California.
** Now at Redford High School, Detroit, Michigan.
[1] R. E. Nusbaum and J. W. Hackett, J. Opt. Soc. Am. **31**, 620 (1941).
[2] R. E. Nusbaum, D. L. Fry, and J. W. Hackett, J. Opt. Soc. Am. **34**, 33–40 (1944).

TABLE II. Synopsis of exposure conditions and associated facts.

Exposure Conditions:
Prespark—30 seconds
Exposure—40 seconds
Spark gap—$\frac{1}{16}\pm0.002''$
Point to plane—one-half the radius from center
Carbons pointed 90° included angle
No air jet
Optical System:
Spectrograph—double Cornu prism instrument constructed in the Laboratory
Slit—Bausch and Lomb, with lens at the slit
Slit width—40μ
Spark electrode position—10 inches from slit
Calibration arc position—26 inches from slit
Region of spectrum focused—2300–2700A
Photographic:
Plate—Eastman Kodak Spectrum Analysis No. 1
Processing—D-19 (4 min. at 69°F), F-5 (2 min. at 69°F), wash and dry

especially, it was necessary to select the short-wave region, i.e., around 2500A. Even in this region there was little choice and this necessitated the selection of non-homogeneous lines; i.e., lines which were considerably different under conditions of spark or arc excitation. The wave-lengths in angstroms of the lines finally selected are tabulated in column 3 of Table I.

In the interest of simplicity and speed only one internal standard line (cobalt) was selected. This proved to be satisfactory since the gamma- of the S.A. No. 1 plate is approximately constant over the wave-length region used.

It should be pointed out that the presence of titanium in Vitallium would have interfered appreciably with the silicon determination as titanium has a strong line superimposed upon the only available silicon line. A careful check for titanium was made at 2527.985 for each sample. Fortunately, none of the samples tested have contained titanium.

The samples were sparked at a point one-half the radius from the center and with the circuit constants listed in Fig. 1. The moving plate method was used to determine the prespark (30 seconds). A 40-second exposure brought the lines to the desired intensity. The reproducibility obtained with these exposure conditions, together with an average of three separate log I ratios, appeared satisfactory for the tolerance of this problem and was used for all determinations.

In order to bring together all the pertinent facts already discussed, a synopsis (Table II) and circuit diagram (Fig. 1) are included.

ANALYTICAL CURVES

The analyses were determined by means of the internal standard method. Every plate was calibrated with a log sector disk.

Proper standards for plotting good analytical curves should provide concentrations bracketing the range which will be found in the unknowns. When the analytical curves were first set up, the samples available for standards did not have enough variation in concentration to determine the slope of the analytical curves. Therefore, the theoretical slope was at first assumed (slope = 1 when logarithms have the same base).

$$\log C = \log(I_x/I_s) + K. \tag{1}$$

Although this is a theoretical curve and seldom obtained in practice, several authors[3–6] have reported analytical curves with this slope and it does give a starting point.

After more chemical analyses were obtained, the curves were rotated to fit the points more exactly.

In most cases where spectrographic methods have been used in the past, the internal standard element has constituted 90 percent or more of the total sample. In these cases there are rarely appreciable changes in the amount of the internal standard material regardless of how the concentrations of minor elements and impurities change.

Most of the Vitallium samples have contained very nearly the same amount of the standard (cobalt). It has therefore been convenient to plot analytical curves in the usual way, assuming that the concentration of cobalt does not change appreciably from one sample to the next.

However, to provide more widely separated points on the analytical curves, two samples, V-189 and V-190, were prepared with a composition somewhat different from the other standards. V-189 contained unusually low concentrations of chromium and molybdenum and V-190 contained high chromium. As a consequence, the former con-

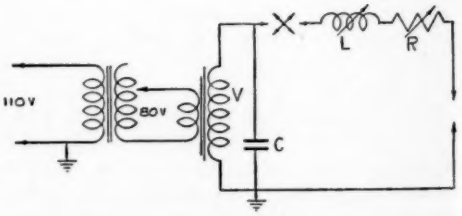

FIG. 1. Spark circuit diagram.

V = 35,000 peak volts
C = 0.021 microfarad
L = None added (0.033 mh residual in the resistance R)
R = 12 ohms

[3] G. A. Nahstoll and F. R. Bryan, J. Opt. Soc. Am. 35, 646 (1945).
[4] M. F. Hasler and H. W. Dietert, J. Opt. Soc. Am. 34, 751 (1944).
[5] M. F. Hasler and J. W. Kemp, J. Opt. Soc. Am. 34, 21 (1944).
[6] M. F. Hasler and H. W. Dietert, J. Opt. Soc. Am. 33, 218 (1943).

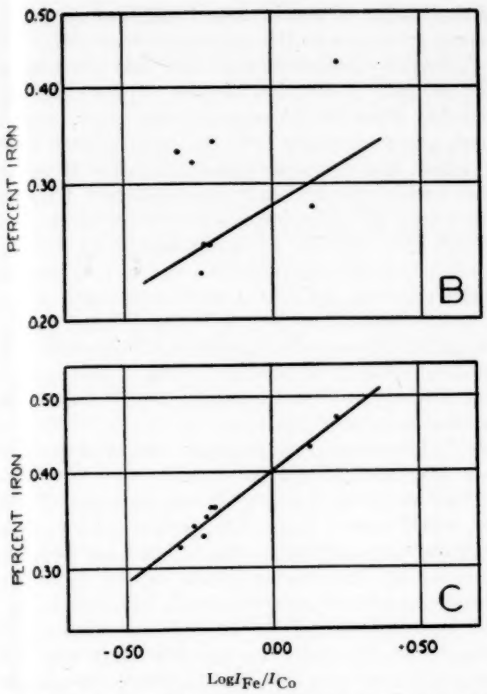

FIG. 2. Two sets of chemical analyses on 8 samples plotted against the same logI_{Fe}/I_{Co} values.

tained 72 percent cobalt and the latter 58 percent. The log I ratios obtained with these samples were corrected for the change in the amount of cobalt, and the new values were plotted against the chemical composition.

REPRODUCIBILITY

In order to examine the reproducibility of the procedures described in this paper, repeat analyses were made on all standard samples. The results of this test are summarized in columns 4 and 5 of Table I. Standard deviations were calculated from the expression

$$\sigma=(\sum D^2/(N-1))^{\frac{1}{2}},$$

where D is the deviation of each determination from the mean for the sample analyzed, and N is the total number of analyses made. For the purposes of this table each determination was obtained from a single exposure. The standard deviations are based on the total number of determinations and are expressions of the reproducibility of the spectrographic analysis.

Because of the rather large deviations shown in some of the analyses of the standard samples (due in part to segregation in the standards), it was the practice in making analyses of unknowns to average three determinations to increase the accuracy. (Standard deviation reduced by $1/\sqrt{3}$.)

AGREEMENT BETWEEN CHEMICAL AND SPECTROGRAPHIC DATA

Some of the samples examined in the course of this work have been analyzed by both the chemical and spectrographic methods. The results of these analyses and the differences between the two methods are shown in column 6 of Table I. This table compares the most recent chemical analyses and the spectrographic analyses as obtained from the most recent analytical curves. Three spectrographic determinations are averaged for the spectrographic value. In general, the agreement between chemical and spectrographic values shown in this table is good. Indeed it is better than the agreement between single spectrographic determinations. It is important here to remember that the spectrographic value reported was the average of three determinations.

It should be pointed out, however, that agreement between the two methods was not nearly this good when the project started. The method was put into operation with a minimum of preliminary development work and then as more samples were submitted and more experience gained, some of the chemical analyses were improved and analytical curves were changed to fit the new values. An example is the iron determination as shown in Fig. 2.

Figure 2 shows the logarithms of line intensity ratios, that is I_{Fe}/I_{Co}, for eight samples plotted against two different sets of chemical analyses in parts B and C. The line shown in each section is the analytical curve which was being used when the chemical analyses were made. In part B, agreement between chemical and spectrographic determinations is not good. However, in part C the

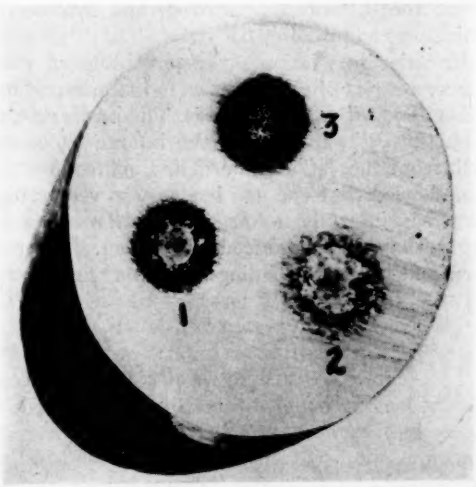

FIG. 3. Effect of gap space on burns. 1. $\frac{1}{32}$". 2. $\frac{1}{16}$". 3. $\frac{3}{32}$".

agreement is most satisfactory. This improvement resulted entirely from increased accuracy in chemical determinations because the same values of $\log I_{Fe}/I_{Co}$ were used throughout. It should be pointed out that the chemical analyses used for part *B* were made as routine controls with a standard method described in the literature, whereas the chemical analyses used for *C* were made by a new method developed after considerable research.

An analytical curve such as shown in Fig. 2, part *B*, would not ordinarily be considered adequate for spectrographic analysis. It was used in this case as a temporary expedient in order to make a contribution to the project without further delay. This contribution could have been made without analytical curves by simply reporting log *I* ratios, if the metallurgists were accustomed to thinking in these terms.

SENSITIVITY OF LINE INTENSITIES TO GAP LENGTH

In the course of this work án important fact became apparent; the gap length had to be very accurately set.

While setting up the procedure, approximately one exposure in 15 was very heavy. This particular effect was the result of an intense burn such as is shown in Fig. 3, burn 3. The counter carbon electrode had been well coated with metal during the exposure. After considerable experimentation this was found to be due to a variation in gap length.

Three exposures were made using a $\frac{1}{32}''$ gap, the normal size gap of $\frac{1}{16}''$, and an oversized gap of $\frac{3}{32}''$. The $\frac{1}{32}''$ gap showed a burn different from the normal ones, but such as would be expected by using a small gap; the $\frac{1}{16}''$ gap produced a normal burn; the $\frac{3}{32}''$ caused the deep and intense burn which was so puzzling.

In order to eliminate the possibility of coincidence, a series of exposures were made using a $\frac{1}{16}''$ gap and a $\frac{3}{32}''$ gap alternately. The burns resulting from the $\frac{3}{32}''$ gap were of the intense type, while each from the $\frac{1}{16}''$ was normal. Further investigation showed that the gap length was very critical, for a spacing only a few thousandths of an inch larger than $\frac{1}{16}''$ produced this effect. Either the burn and its corresponding spectrum were normal, or they were of the intense type, and no gradual change from one to the other was observed.

A further study showed that this phenomenon was partially a function of time. If, for example, a sparking time of 70 seconds was used, and the spark gap was greater than $\frac{1}{16}''$, a deep burn always occurred. If the sample was sparked for 40 seconds with a wide gap, no heavy burn resulted. However, when two 40 second exposures were made, even though the area for each exposure was different, the heavy burn occurred with the second exposure if the gap was $\frac{3}{32}''$. In a similar manner, two 30 second exposures with a $\frac{3}{32}''$ gap did not produce the effect but the third one did (allowing only a few seconds between exposures). When the sparking time was reduced to 20 seconds, repeated exposures resulted in normal burns.

This interesting phenomenon was later found to be unique with Vitallium alloy.

Very recently an article has appeared in the literature[7] which may cast some light on this problem. The authors of this article have found the same effect in point-to-plane sparking of iron alloys when low power (0.007μf, 35,000v) is used, and have correlated the dark burns with sample temperature. It may be possible that the same thing will hold true for Vitallium with the sparking conditions given in this paper (0.021μf, 35,000v).

DISCUSSION OF RESULTS

A method of analyzing Vitallium alloys with a spectrograph has been developed. The method has been extended to a number of high temperature alloys with considerable success. Even while setting up the procedure without adequate standards, it was possible to give analytical results which were useful in control of melting processes.

A spectrographic method having good reproducibility can be used to control most processes, even though exact chemical values are not obtained. It would often be sufficient to report log *I* ratios. However, due to the fact that log *I* ratios are frequently not understood, it is possible to give results in terms of percent composition which are accurate enough in relative value to furnish a control, even though the absolute compositions are not correct.

[7] D. M. de Waal and S. M. Naudé, Spectrochimica Acta III, 127–40 (1948).

Optical and Electrical Alterations to a Michigan Microphotometer*

H. H. GROSSMAN, E. W. PETERSON, J. L. SAUNDERSON, AND V. J. CALDECOURT
The Dow Chemical Company, Midland, Michigan

(Received December 13, 1948)

A microphotometer of the type described by Vincent and Sawyer has been altered optically and electrically to eliminate certain objectionable features. Scattered light, dependence of the plate projection lens upon the aperture of the plate illumination system, photo-cell fatigue, relatively short life of the projection lamp, and the need for a constant voltage power supply have been eliminated. Plate calibration curves obtained with the altered instrument on wide bands and narrow lines are essentially identical.

A MICROPHOTOMETER of the type described in 1941 by Vincent and Sawyer[1] and constructed in the University of Michigan Physics Department shops has been in almost constant use in this laboratory since it was obtained in 1942. Whereas the mechanical design of the instrument has proved to be almost ideal with essentially no evidence thus far of mechanical wear, several features of the optical and electrical system have caused difficulty.

The original optical system of the microphotometer can be followed with reference to Fig. 2 of the Vincent and Sawyer paper. In the original system, the closely wound helical filament of the lamp is projected on the photographic plate at approximately unit magnification, and at the same time "out-of-focus" light from the same lamp illuminates a larger area of the plate for viewing the spectrum being measured. The plate thus illuminated is in turn imaged at about tenfold magnification on a screen containing a slit behind which is mounted the photo-cell. The intense filament image enlarged to a width of about $\frac{1}{4}$ inch is centered about the photo-cell slit, the latter being somewhat narrower than the projected width of lines to be measured. A variable aperture on the first lens used for projecting the filament onto the plate is provided for adjusting the clear plate reading. Another optical train projects a reference spectrum onto the screen to aid the operator in locating lines.

The undesirable features of our instrument have been chiefly, (1) the presence of an important amount of scattered light, (2) dependence of the position of best focus of the plate projection system upon the aperture of plate illumination system, (3) dependence of the maximum blackness reading of a spectral line on the aperture of the plate illumination system and the width of the line, (4) relatively short life of the illuminating lamp, and (5) fatigue of the barrier layer photo-cell. From the operators' standpoint, the wide intense filament image around the slit is also objectionable because weak spectral lines which are visible by the general plate illumination are completely lost when brought into the very bright area close to the slit.

The difficulties connected with the variable aperture of the plate illumination system could of course be avoided by working at fixed aperture and using a variable shunt on the galvanometer for adjusting the clear plate reading. However, the remaining undesirable features appeared to justify the rather extensive optical and electrical alterations which were undertaken.

In a microphotometer of the original design, in which the intensely illuminated area of the plate is considerably larger than the line to be measured, it is imperative that the focus of the plate projecting lens be excellent and that the amount of light scattered in this lens be negligible. If these conditions cannot be met in practice, the simplest alternative is to change the plate illumination system in an attempt to confine the light passing through the plate to an area smaller than that of the line to be measured. This can be accomplished by projecting the image of an illuminated slit onto the plate but it is obviously as difficult to confine all the light to an area within the line as it is to obtain a perfect plate projection system which excludes all of the light from the photo-cell slit except that which passes through the line. However, a double slit arrangement of this sort has a great advantage over the original single slit arrangement inasmuch as any unwanted light contributing to the photo-cell current must be scattered twice, once in the plate illumination system and again in the plate projection system. The practical requirements on focus and scattered light are thus not nearly so exacting in either lens of the double slit arrangement as in the plate projection lens of the single slit arrangement.

Briefly, the alterations included projection of a slit narrower than the line to be measured on the plate instead of the broad image of the lamp filament as used formerly, provision of filtered light for general plate illumination and a second filter

* Presented at the Detroit meeting of the Optical Society of America, October 21–23, 1948.

[1] H. B. Vincent and R. A. Sawyer, J. Opt. Soc. Am. **31**, 639 (1941).

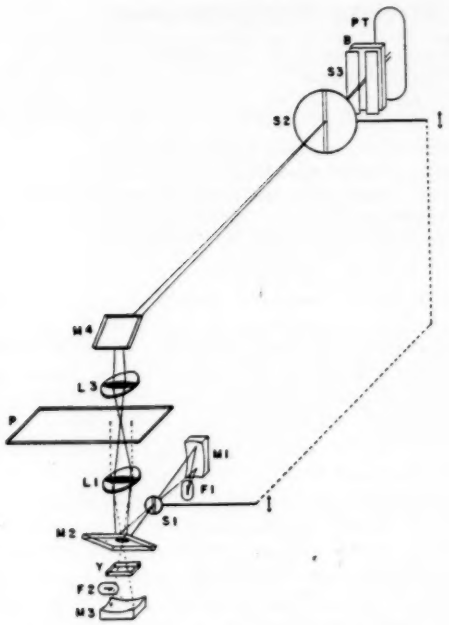

FIG. 1. Schematic of the revised optical system.

to keep this light from the light sensitive element, substitution of an electron multiplier photo-tube for the barrier layer cell, and development of an electrical circuit which eliminates the need for a light source of constant intensity.

The optical system of the revised microphotometer is shown schematically in Fig. 1. Light from a filament F_1 is condensed by the spherical mirror M_1 upon the slit S_1, and the slit S_1 is imaged on the plate P by the lens L_1 after reflection from the plane mirror M_2. The lamp is the same type as used in the original microphotometer, but operated at 6 volts instead of the rated 8.5 volts. The focusing lens L_1 is the same as the first lens L_1 in the original instrument, and the original lens L_2 is not used. The slit S_1 is about 25μ wide and is projected onto the plate at the same width. The plane mirror M_2 is a focal length away from the lens L_1, and has a rectangular aperture about 1×3 mm where the aluminized surface has been removed. The light from the second lamp F_2 is condensed upon the aperture in the mirror after passing through the yellow filter Y. This yellow light is rendered parallel by lens L_1 and serves for the general plate illumination. The slit S_1 is mounted in a bearing for rotation and is linked to the slit S_2 in front of the phototube so that both slits rotate together. Each of the slits is rotated by turning a shaft which operates a screw drive. The two shafts are linked by means of sprockets and a chain.

The plate projection system was not changed, except that a 931A multiplier photo-tube was substituted for the barrier layer photo-cell. A blue filter B, ground on one side for good light scattering properties, is mounted with a fairly large slit S_3 in front of the multiplier tube. The slit S_3 makes the instrument quite insensitive to room illumination, and the blue filter prevents the fraction of yellow light which is scattered in the plate projection system from reaching the photo-tube. The slit S_2 is 500μ wide so that the image of the 25μ slit S_1 when enlarged to 250μ on the screen always stays well within the slit S_2. The width of slit S_2 is relatively unimportant, since, with a perfect plate illumination system (lens L_1) in which no white light passed around the line on the plate, the only function of slit S_2 would be to exclude room illumination. However, with an imperfect plate illumination system, the full advantage of the double slit system is utilized only if S_2 is narrower than the projected image of the narrowest line being measured. This criterion assures that unwanted light will have to be scattered twice.

The original lamp was left in its original mounting to serve for projecting the reference spectrum. Except for the lens L_1, which was remounted with its axis vertical, the other optical parts were removed from this compartment. The additional space required for the revised optical system was obtained by attaching a box to the bottom of the original compartment. This box projects about 6 inches below the top of the table on which the instrument is mounted.

The electronic circuit is shown in Fig. 2. The 931A on the left monitors the lamp, and, through the 6J7 control tube, controls the voltage applied to itself in such a way that its plate current tends to remain constant.[2] The second 931A which measures the photometer beam has the same voltage applied to its electrodes, and since the voltage-gain characteristics are nearly alike, its plate current changes very little with lamp fluctuations.[3] The lamps are supplied with a.c. and the entire unit is supplied from a Sola regulator. The measuring photo-tube drives the original galvanometer through the amplifier shown. Since the galvanometer requires several microamperes of current, the amplifier is necessary to avoid fatigue of the photo-multiplier. With this circuit the photo-current requirement is about 0.2 microampere and fatigue is not appreciable.

The amplifier[4] has 100 percent inverse feedback so its gain is nearly unity. The output from the 12BE6 feeds back through the 6C4 to the grid resistor so that the grid 3-to-cathode potential tends to remain constant. The galvanometer meas-

[2] M. H. Sweet, J. Opt. Soc. Am. **36**, 432 (1946).
[3] Subject of patent proceedings.
[4] H. S. Anker, Electronics, June, 1947, p. 138.

ures the voltage fed back to compensate for the photo-current. The filament voltage on the 12BE6 is set at 8 volts to keep grid current low. The large cathode resistor of the 12BE6, and the first grid reduce fluctuations caused by changes in filament temperature. The potential of grid 1 is adjusted so that the galvanometer current is zero when the current through the input resistor is zero. The clear plate reading is adjusted by means of an Ayrton shunt on the galvanometer, not shown.

In order to check the performance of the revised instrument and to demonstrate the difficulties encountered before revision, a test plate was made. On this plate are two sets of lines and two sets of bands, so that calibration curves may be drawn for both lines and bands. One set of lines and one set of bands received identical exposures, while the other two sets differed from these by a filter factor.

This plate was made by exposing two areas on the plate simultaneously. These areas were defined by two slits, 100μ wide and $\frac{1}{8}$-inch wide, placed very close to the plate, and illuminated from a small symmetrically placed source. Successive exposures were made by moving the plate under the slits, and the intensity was varied by changing the distance to the light source. The exposure time was constant. The exposures through the filter give data to check the self-consistency of the method.

The plate was read under three microphotometer conditions, and the results are shown in Fig. 3. The closed figures represent exposures through the filters. That the closed and open figures lie on the same curves shows that the calculated log relative intensities are correct and that the plate is very uniform. It should be pointed out that the open and closed figures that occur close together are not corresponding points. These points are close just because the filter factor happens to be very nearly equal to two exposure steps.

All readings were taken after the slit in front of the photo-tube had been increased from 250μ to 500μ, but since the line width was 100μ the results are indicative of the regularly used condition of a 50μ line with a 250μ slit.

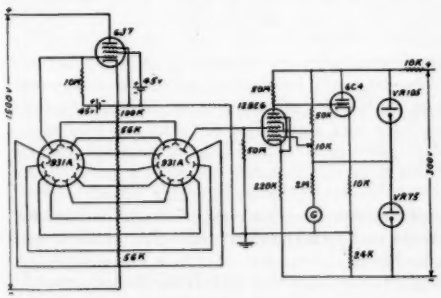

FIG. 2. Electrical system showing the voltage-regulating and lamp monitoring device and the impedance matching amplifier between the measuring photomultiplier and the galvanometer.

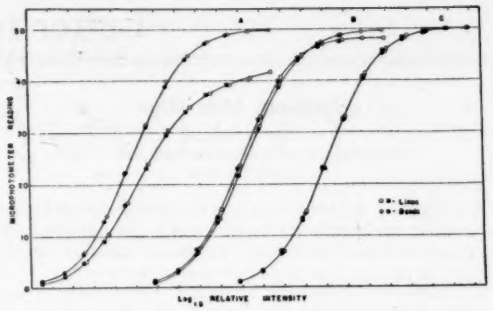

FIG. 3. Calibration curves for a photographic plate obtained from readings on lines and bands, taken with different microphotometer conditions: (A) Original microphotometer with full aperture; (B) Original microphotometer with small aperture; (C) Revised microphotometer.

The A curves were obtained from readings taken before the general revision, and with the aperture fully open. The large difference between the curves for lines and bands demonstrates the presence of a large amount of scattered light getting into the slit from the region close to the line.

The B curves were obtained from readings taken before the general revision, with the aperture as nearly closed as it could be with the lamp operated at 8.25 volts. For dark lines the effect of scattered light is still easily seen. We have no adequate explanation of why these curves cross over, so that light lines appear to be darker than light bands.

The C curve was obtained from readings on the revised instrument. The curves for lines and bands merge almost completely except for a region near the knee of the curve where the lines appear to be slightly darker than the bands. This is probably due to the Eberhard effect.

Before the instrument was revised, since it was used principally on magnesium alloy analysis, several different plate calibration curves were necessary for best results. The reason for this is that the lines obtained from magnesium alloys vary considerably in sharpness and background and often a diffuse magnesium line is compared to a sharper line of another element. However, if a single plate calibration curve which was a sort of average curve were used appreciable analytical error resulted, especially when using the variable aperture to adjust the clear plate reading. The revised instrument has eliminated this difficulty, since the shape of the spectral lines no longer has an important effect upon the plate calibration curve.

Letter to the Editor

Optimum Absorption

JOHN E. TYLER
Interchemical Corporation, New York, New York
January 3, 1949

IN absorption spectroscopy, it is frequently desirable to determine the sample thickness which on analysis will show the greatest separation between two points at selected wave-lengths on the curve of transmittance *vs.* wave-length.

If a transmission-density conversion table and a spectrophotometric transmittance curve of the material in some known thickness are available, this optimum thickness can be readily calculated by means of the equation—

$$a = w \log \frac{D_2}{D_1} \Big/ (D_2 - D_1),$$

in which,

a is the optimum thickness,
w is the sample thickness for the known curve,
D_1 and D_2 are the densities on the known curve at the selected wave-lengths.

This equation is derived from the exponential law and can be applied to the case of varying concentrations if interfering reactions are absent. Its accuracy falls off where the transmittances involved are too low or too high to be accurately determined.

Book Review

Vision and the Eye

M. H. PIRENNE. The Pilot Press, Ltd., London, 1948. pp. xx+187.

In this small book, Dr. Pirenne has attempted to describe for the "interested but non-specialist reader," a few of the fundamental physiological properties of the eye. He has drawn his information from the results of experiments performed on eyes of different organisms, including those of humans, insects, *Limulus polyphemus*, attempting thereby to integrate some of the fundamental knowledge of visual physiology.

Dr. Pirenne's approach is entirely physiological, whereby he considers the eye as a physical mechanism which can be explained in terms of matter and energy. He does not consider the problems of sensation and consciousness, although in his very brief last chapter, "Physics and the Phenomena of Life" he discusses the relationship between the two approaches. Unfortunately, he contributes little to an understanding of the philosophical problems involved in these considerations.

Wisely enough, Pirenne has chosen to discuss a few selected topics. Almost entirely omitted are the problems of brightness discrimination and dark adaptation, among others. The topics which are discussed are given varying degrees of emphasis, as they should be. However, the over-all presentation leads to confusion about the author's notion of the "non-specialist." For example, in his first chapter "The Eye and the Formation of the Retinal Image," the exposition is entirely elementary. That portion concerned with geometrical optics and formation of the retinal image is superfluous for any student even with an elementary knowledge of these subjects; on the other hand, it is far too sketchy for someone without such an elementary knowledge. Again, the chapter, "The Two Eyes and the Brain," is lacking, perhaps, in enough details, so that even a more advanced reader could not receive a clear glimpse of the problem.

Dr. Pirenne has discussed several topics with success, notably the nervous activity produced in the eye by the action of light, the structure of the retina, and the properties of rods and cones, and visual acuity of the eyes of insects and of man. In the latter instance, he has shown the value and interest of studies in comparative physiology.

The subject which is discussed in greatest detail and with greatest clarity concerns the recent researches into the quantum nature of light and its relationships to the retinal mechanism and theory of vision. Pirenne, who was concerned in the pioneer investigations of these problems in the laboratory of the late Dr. Hecht, draws upon his first-hand experience to present these recent and significant concepts with simplicity and understanding.

As intimated above, it seems to this reviewer, that in the discussions of other subjects, the author has not succeeded in his attempts. This might be illustrated best in relation to the sections on color vision. Thirty pages are devoted to the problems of color vision. This is a great deal of space compared to the pages devoted to most other topics but it is not enough in which to discuss, even from an elementary point of view, Newton's doctrine of color, normal and abnormal color vision, and Thomas Young's theory. In trying to cover so great a field in so few pages, the exposition suffers. And since the author is discussing, presumably, the latest researches, it should be mentioned that although Young's theory of color vision, containing elements of straightforwardness and simplicity, has led to increased knowledge and clearness regarding ideas of color vision, the modern picture of this field is not complete without mention of more recent work extending and modifying even some basic concepts based on the trireceptor theories.

Although it is difficult to see how that group of "non-specialist readers" who possess little or no knowledge of the eye and of the physiology of vision, could receive from this book an entirely clear picture of many of the visual functions discussed, it should be said that the straightforward style of presentation plus the selected bibliography at the end of each chapter would be of great help to the interested reader. Despite the brevity of presentation, Pirenne does succeed in conveying a good impression of the challenging nature of the problems concerned with vision. He presents his topics from the point of view of research, giving as a basis for some of his discussions (notably those concerning visual acuity and the quantum story) simplified experimental conditions. These factors do tend to make the book useful for the intended purpose, and it is unfortunate that it is not entirely successful. Writing for a non-specialist (whatever his background may be) has always been a difficult and challenging task. In many fields of research, including that of the physiology of vision, writing for the non-specialist has been attempted too infrequently. Books, such as the one which Dr. Pirenne has written, should be welcomed and encouraged.

IRVING H. WAGMAN
Jefferson Medical College, Philadelphia